

ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ

31
П-28

Вини

Лавренко В. 474. 22

ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ

Издание третье, переработанное и дополненное

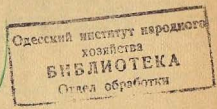
Допущено Министерством высшего и среднего
специального образования СССР в качестве
учебника для студентов экономических
специальностей вузов



2009
ПРОВЕРКА



387010



Москва «Статистика» 1975

В. С. КОЗЛОВ, Я. М. ЭРЛИХ,
Ф. Г. ДОЛГУШЕВСКИЙ, П. И. ПОЛУШИН

О-28 **Общая теория статистики. Учебник для студентов экон. специальностей вузов. Изд. 3-е, перераб. и доп. М., «Статистика», 1975.**

392 с. с ил.

На обороте тит. л. авт.: В. С. Козлов, Я. М. Эрлих, Ф. Г. Долгушевский, П. И. Полушин.

В учебнике отражены изменения в теории и практике статистического учета, которые произошли после выхода в свет предыдущего издания. Это, в частности, усовершенствование статистической отчетности, широкое внедрение электронно-вычислительной техники и др. Пересмотрены и обновлены примеры, иллюстрирующие те или иные теоретические положения.

10805—076
О 008 (01) — 75 105—75

© Издательство «Статистика», 1975.

ОТ АВТОРОВ

Предыдущее издание учебника вышло в 1967 г. Решения XXIV съезда КПСС, задача создания Автоматизированной системы государственной статистики, новое Положение о ЦСУ СССР, а также все более широкое применение в практике экономической работы методов математической статистики — все это вызвало необходимость существенных изменений и дополнений в учебнике.

В данном издании учебника приняты во внимание и учтены многие замечания и пожелания, которые были высказаны в рецензиях на второе и рукопись третьего издания, в письмах читателей, на научно-методических конференциях и семинарах. Изучен и обобщен также опыт преподавания курса.

Большая часть глав в третьем издании значительно переработана как по содержанию, так и по структуре. Изложены многие вопросы, которые не рассматривались во втором издании. В I главе больше внимания уделено категориям статистической науки. В III и IV главах освещен вопрос о сборе и обработке информации с помощью электронных вычислительных машин. Глава VI дополнена параграфом «Анализ вариационных рядов», где рассмотрены вопросы о закономерностях и кривых распределения, критериях согласия, показателях асимметрии и эксцесса.

В главе VIII расширена классификация связей: дано понятие о стохастической связи, частным случаем которой является корреляционная связь. Различные методы изучения взаимосвязей рассматриваются последовательно один за другим, а метод анализа регрессии и корреляций излагается по его этапам. Больше внимания уделено логическому содержанию и статистическому смыслу показателей регрессии и корреляции. Показана методика проверки их значимости (существенности). Дано представление о возможностях использования регрессионно-корреляционных моделей для выявления резервов.

В главе IX более полно изложены вопросы анализа рядов динамики, в частности, в тех случаях, когда уровни ряда (или цепные абсолютные приросты) сильно колеблются. Наряду с выравниванием по прямой, рассматривается выравнивание ряда дина-

мики по показательной кривой и параболе второго порядка, а также смысл параметров этих кривых. Более обстоятельно изложен вопрос об интерполяции и экстраполяции динамических рядов.

В главе X дополнительно рассмотрен вопрос о построении и исчислении индексов в условиях несравнимого ассортимента (круга единиц). Наряду с цепным методом построения системы взаимосвязанных индексов и разложения прироста по факторам рассмотрены и другие возможные пути решения этой задачи. Кроме случая двух факторов — сомножителей, рассматривается также случай трех и большего числа факторов.

Книга написана авторским коллективом Одесского института народного хозяйства в составе: доцент В. С. Козлов (руководитель коллектива и научный редактор) — гл. IX, X, п. 4 гл. IV п. 4, гл. V; доцент Я. М. Эрлих — гл. I, IV (кроме п. 4), VII; старший преподаватель П. И. Полушин — пп. 1—3 гл. V. Главы VI и VIII написаны совместно В. С. Козловым и Я. М. Эрлихом; гл. II и III написаны профессором Ф. Г. Долгушевским.

Авторы приносят благодарность коллективу кафедры общей теории статистики Киевского института народного хозяйства за ценные критические замечания и рекомендации, а также выражают признательность всем товарищам, высказавшим замечания по второму изданию учебника.

Отзывы и пожелания по улучшению учебника просим присылать по адресу: г. Одесса, ул. Советской Армии, 8, Одесский институт народного хозяйства, кафедра статистики.

1. ВОЗНИКНОВЕНИЕ И РАЗВИТИЕ СТАТИСТИКИ

Возникновение статистики

Термин «статистика» происходит от латинского слова «status» (статус), что означает состояние, положение вещей, событий.

От корня этого слова образовалось итальянское слово «stato» (стато) — государство, управляемая область. Лич. обладающих знаниями об устройстве и состоянии дел в различных государствах, т. е. государственных деятелей, политиков, называли «statista» (статиста). От этого же корня образовалось и существительное «statistica» (статистика).

В настоящее время с термином «статистика» связан ряд различных и далеко не одинаковых представлений. Так, под статистикой понимают практическую деятельность по собиранию и обобщению данных о различных явлениях и процессах общественной жизни. Например, на промышленных предприятиях собирают данные о выпуске продукции, явках и неявках работающих и т. д.

Под статистикой понимают, далее, совокупность числовых показателей, характеризующих общественные явления и процессы. В этом смысле говорят о статистике производства, статистике производительности труда, статистике заработной платы и т. д. Наконец, под статистикой понимают особую отрасль общественных наук. Об этом и будет идти речь в учебнике.

Статистика как общественная наука возникла из практических потребностей общественной жизни. Зарождение статистической науки связано с обобщением опыта различных работ учетно-статистического характера. Потребность в проведении таких работ, как свидетельствует история, появилась еще в глубокой древности.

С образованием классового общества и государства возникает потребность в сведениях о численности населения. Для получения этих данных производился учет, в первую очередь учет численности взрослого мужского населения, пригодного к военному делу.

Кроме сведений о численности населения, государство интересовалось также его податной способностью. С этой целью производились описи и оценка имущества, земли, скота и др.

В дальнейшем, по мере развития производительных сил и соответствующих им производственных отношений, значительно расширяется круг явлений, сведения о которых необходимы для нужд общественной жизни. Вслед за населением и его имуществом учетно-статистические работы охватывают торговлю, промышленность, сельское хозяйство, финансы. В то же время эти работы начинают проводиться все более регулярно.

Регулярное проведение статистических операций потребовало теоретического осмысливания всей работы по сборанию, обработке и анализу статистических данных, что и вызвало к жизни новую отрасль науки об обществе — статистику.

Зарождение статистической теории часто связывают с именем Вильяма Петти (1623—1687), которого Карл Маркс назвал отцом буржуазной политической экономии и в некотором роде изобретателем статистики. Заслуга В. Петти в том, что он впервые применил численный метод для анализа закономерностей общественной жизни. Свою работу, где он стал говорить «языком чисел», В. Петти назвал «Политическая арифметика» (1676 г.).

Говоря о статистике как науке, следует подчеркнуть ее исторический характер, т. е. тот факт, что возникновение, развитие и содержание статистики связаны с развитием общественных производительных сил и характером производственных отношений.

В капиталистическом обществе объем статистических работ по сравнению с эпохой феодализма значительно увеличивается.

Это связано с обобществлением производства и ростом потребностей буржуазного государства и отдельных капиталистов в статистических данных о рынках сырья и сбыта продукции, сферах приложения капитала и т. д.

Возникают специальные статистические органы, занимающиеся систематическим сборанием и обработкой статистических данных по различным вопросам общественной жизни. Вслед за национальными появляются *международные* статистические органы: различные статистические институты, бюро и т. д. Начинают также проводиться *национальные*, а затем и *международные статистические съезды, конгрессы, конференции*.

Развитие статистической практики вызывает интерес к статистическим знаниям. Статистику начинают преподавать в высших и средних учебных заведениях. Все в большем количестве издаются работы по теоретическим и практическим вопросам статистики. Наряду с монографиями в ряде стран начинают выпускать специальные статистические журналы.

Вместе с расширением поля деятельности статистики, возрастанием числа статистических исследований разнообразных проявлений жизни общества растет и число ученых-статистиков, равняющихся статистическую теорию.

Из зарубежных статистиков особо следует выделить бельгийского ученого Адольфа Кетле (1796—1874). Большая заслуга Кетле в том, что он, по словам К. Маркса, доказал, что даже

возникающая случайности общественной жизни обладают внутренней необходимостью. Однако объяснить эту необходимость Кетле никак не удалось.

Значительное развитие получает статистическая теория в трудах русских ученых. Это в первую очередь академик К. Ф. Герман (1766—1815), автор первого русского оригинального труда по теории статистики «Всеобщая теория статистики», вышедшей в 1809 г. Затем следует плеяда таких талантливых ученых, как Д. П. Журавский (1810—1856), П. П. Семенов-Тянь-Шанский (1827—1914), Ю. Э. Янсон (1835—1893), А. И. Чупров (1842—1908), А. А. Чупров (1874—1926), А. А. Кауфман (1864—1919) и многие другие.

Большую роль в развитии статистической теории сыграла и русская земская статистика, которую, несмотря на ряд крупных недостатков, высоко ценил В. И. Ленин.

В эпоху капитализма в статистике, как и в других общественных науках, наряду с прогрессивными, получают развитие и различные вульгарные школы и течения, ненаучные и реакционные по своему существу, ставящие своей целью апологетику капитализма. В условиях капиталистического общества буржуазная статистика, как и буржуазная экономическая наука в целом, вырождается и становится средством обмана трудящихся масс и защиты интересов буржуазии.

Классовую направленность буржуазной экономической науки вообще и статистики, в частности, неоднократно разоблачали классики марксизма-ленинизма. Еще Маркс указывал, что с момента завоевания буржуазией политической власти наступает смертельный час для научной буржуазной политической экономии, так как отныне дело идет уже не о том, правильная или неправильная та или другая теорема, а о том, полезна она для капитала или вредна. «Бескорыстное исследование уступает место сражениям наемных писак, беспристрастные научные изыскания заменяются предвзятой, угодной апологетикой»¹.

В. И. Ленин в ряде произведений подвергает уничтожающей критике официальную буржуазную статистику за фальсификацию данных, апологетику капитализма, неумение и нежелание научно обработать богатый фактический материал. «Интересы буржуазии», — указывал В. И. Ленин, — требуют прикрашивания капитализма и затуманивания классовых пропасти»².

Лживость официальной буржуазной статистики с пиничной открытостью признают сами представители буржуазии. В этом отношении характерно высказывание известного английского государственного деятеля второй половины XIX в. Дизраэли о том, что существуют три вида лжи: просто ложь, злая ложь и статистика. Со времен Дизраэли положение не изменилось к лучшему, наоборот, лживость, тенденциозность буржуазной статистики, особенно официальной, достигла небывалых размеров.

¹ Маркс К. и Энгельс Ф. Соч., т. 23, с. 17.

² Ленин В. И. Полн. собр. соч., т. 23, с. 435.

В капиталистических странах фальсификации подвергаются данные о безработице, средней заработной плате, потреблении населения и т. д. Это необходимо иметь в виду при использовании публикаций буржуазной статистики.

Возникновение и развитие марксистско-ленинской статистики

В эпоху капитализма, наряду с буржуазной статистикой, зарождается и получает значительное развитие и подлинно научная статистика — статистика рабочего класса. Ее создание связано с деятельностью Интернационала, в уставе которого был специальный пункт о статистике труда.

Основоположниками подлинно научной статистики являются Карл Маркс и Фридрих Энгельс. Их политико-экономические работы, особенно «Капитал» К. Маркса и «Положение рабочего класса в Англии» Ф. Энгельса, не только насыщены новейшими для того времени фактическими данными о характере капиталистического производства и положении рабочего класса при капитализме, но и содержат ряд высказываний по важнейшим вопросам статистической теории.

Дальнейшее развитие подлинно научная статистика получает в трудах В. И. Ленина. Отмечая, что коренные вопросы развития экономики нельзя разрабатывать без учета массовых данных, собранных по одной определенной программе и сведенных вместе специалистами-статистиками, В. И. Ленин говорил, что «социально-экономическая статистика — одно из самых могущественных орудий социального познания»¹. Именно поэтому, считал В. И. Ленин, рабочему классу нужна своя статистика, чтобы лучше познать свое собственное движение и обеспечить этим его успехи.

С именем Ленина связан новый, вышедший этап в развитии статистической науки. В трудах В. И. Ленина даны многочисленные непревзойденные образцы политико-экономического анализа материалов русской и зарубежной статистики. Критически разрабатывая статистические данные, В. И. Ленин формулирует важнейшие принципы статистической методологии, которые легли в основу советской статистики.

Уже после Великой Октябрьской социалистической революции В. И. Ленин много внимания уделял развитию советской статистики. Отмечая принципиальное отличие советской статистики от статистики капиталистической, Ленин писал: «Статистика была в капиталистическом обществе предметом исключительного ведения «казенных людей» или узких специалистов, — мы должны понести ее в массы, популяризировать ее, чтобы трудящиеся постепенно учились сами понимать и видеть, как и сколько надо работать, как и сколько можно отдыхать...»².

За годы Советской власти советская статистика стала одной

из самых популярных наук. Охватывая все проявления общественной жизни, она становится необходимым подспорьем в грандиозной работе по строительству коммунизма в нашей стране.

Свое дальнейшее развитие советская статистика получила в решениях съездов КПСС, постановлениях ЦК партии и Советского правительства, а также в трудах советских ученых — академиков С. Г. Струмилина, В. С. Немчинова, членов-корреспондентов АН СССР В. Н. Старовского, М. В. Птухи, докторов экономических наук А. Я. Боярского, А. И. Гозулова, А. И. Ежова, А. И. Ротштейна, Б. С. Ястремского и многих других.

2. ПРЕДМЕТ СТАТИСТИКИ

Объект и предмет статистики

Содержание каждой самостоятельной науки определяется в первую очередь ее предметом, т. е. тем, что изучает данная наука. Статистика как общественная наука имеет объектом своего изучения общество, явления и процессы общественной жизни.

Но общество, общественные явления изучает не только статистика; общество является объектом изучения всех общественных наук. При этом каждая из них изучает одну из граней общественной жизни, которая выступает как предмет данной науки. Что же изучает статистика в явлениях и процессах общественной жизни?

Предметом статистики являются размеры и количественные соотношения массовых общественных явлений, закономерности их течения и развития.

Рассмотрим теперь подробнее составные элементы понятия о предмете статистики.

Количественная сторона общественных явлений

Первая отличительная особенность предмета статистики заключается в том, что статистика изучает количественную сторону общественных явлений и процессов в конкретных условиях места и времени.

Говоря о количественной стороне общественной жизни, имеют в виду прежде всего конкретные размеры общественных явлений и процессов. Например, статистика изучает численность населения и процессов. Например, статистика изучает численность населения земного шара, распределение населения по континентам, государствам и т. д. на ту или иную дату. Так, численность населения земного шара сейчас приближается к 4 млрд. человек, а численность населения СССР на 9 августа 1973 г. составляла 250 млн. человек.

Статистика изучает конкретные размеры общественного производства, распределения, потребления, явлений культурной жизни, что позволяет ей характеризовать экономическую мощь государства, уровень материального благосостояния населения и т. д.

Другим выражением количественной стороны общественной жизни являются числовые соотношения размеров общественных

¹ Ленин В. И. Полн. собр. соч., т. 19, с. 334.

² Там же, т. 35, с. 192.

явлений. Так, статистика установила, что в 1973 г. в СССР в расчете на каждую 1000 человек населения число родившихся составило 17,6, а число умерших — 8,7, откуда естественный прирост — 8,9. В 1974 г. объем промышленного производства в СССР по сравнению с 1973 г. увеличился на 8,0%. В 1973 г. в нашей стране в расчете на душу населения произведено: электроэнергии — 3662 кВт·ч хлопчатобумажных тканей — 26 м², молока — 354 кг и т. д.

Изучение числовых соотношений размеров общественных явлений позволяет статистике характеризовать ход выполнения народнохозяйственных планов, динамику размеров общественного производства, обеспеченность населения различными продуктами и т. д.

Следует подчеркнуть, что изучение количественной стороны общественных явлений, как того требует диалектика, неразрывно связано с учетом так называемых *качественных особенностей* исследуемых объектов, т. е. их *социально-экономического содержания*, так как количественная и качественная стороны общественных явлений взаимосвязаны и обусловлены. Например, исследование объема промышленного производства или национального дохода требует предварительного определения самих понятий «промышленная продукция» и «национальный доход». Чтобы подсчитать численность грамотного населения, необходимо в качестве предпосылки иметь определение понятия «грамотность», т. е. кого, при каких условиях считать грамотным и т. д.

Явления общественной жизни, изучаемые статистикой, находятся в непрерывном изменении и развитии. С течением времени изменяются как их размеры, так и различные соотношения между ними. Неодинаковы они и на разных территориях — в отдельных странах, районах и т. п. Поэтому количественную сторону общественных явлений статистика всегда изучает в конкретных условиях места и времени.

Массовые явления общественной жизни Вторая особенность предмета статистики заключается в том, что статистика изучает количественную сторону массовых явлений общественной жизни, т. е. совокупностей, состоящих из большого числа отдельных единиц. Поэтому конкретные размеры и числовые соотношения размеров общественных явлений выступают как *обобщающие статистические показатели*.

Приведенные выше показатели численности населения СССР, рождаемости, смертности, естественного прироста и т. д. являются обобщающими статистическими показателями, так как выражают в конкретных условиях места и времени закономерности такого массового общественного явления, как воспроизводство населения.

В работе «Статистика и социология» В. И. Ленин подверг уничтожающей критике часто применяемый буржуазными экономистами и социологами прием «выхватывания отдельных фактов»,

так как подобрать отдельные примеры в подтверждение той или иной мысли не так уж трудно, но такое аргументирование не обладает силой убедительности. Исследователь должен, как указывал В. И. Ленин, брать для исследования «не отдельные факты, а всю совокупность относящихся к рассматриваемому вопросу фактов, без единого исключения»¹.

Так, например, при изучении связи между производительностью труда и квалификацией рабочих промышленного предприятия исследователь может обнаружить такие случаи, когда с повышением квалификации производительность труда рабочих растет, и такие случаи, когда с ростом квалификации она падает. Следовательно, на основе отдельных фактов можно прийти к диаметрально противоположным выводам. Но такого рода умозаключения, составленные на основе выхватывания отдельных фактов, не имеют общего с действительно научным изучением вопроса не имеют. Только взяв всю сумму фактов, подвергнув обследованию всех рабочих предприятия, можно получить научно значимый вывод о влиянии квалификации на производительность труда.

Статистическая совокупность Третья особенность предмета статистики заключается в том, что изучаемые статистикой совокупности общественных явлений имеют две характерные черты: 1) *качественная однородность единиц* и 2) *вариирование изучаемых признаков*. Совокупности, обладающие этими чертами, называются *статистическими*.

Важнейшая черта статистической совокупности заключается в том, что она состоит из таких единиц (элементов), которые являются *однородными*, т. е. однородными в определенном отношении; они имеют определенное единство, обладают общими свойствами, объединены общим законом развития. Например, при изучении сельскохозяйственного производства в СССР из общего числа сельскохозяйственных предприятий прежде всего должны быть выделены две статистические совокупности: 1) колхозы и 2) совхозы. В этом случае отдельные колхозы и совхозы выступают как единицы соответствующей статистической совокупности.

Единицы изучаемой статистической совокупности представляют собой качественно однородные первичные элементы этой совокупности. Так, качественная однородность совхозов заключается в том, что это государственные сельскохозяйственные предприятия, у которых все средства производства и производственная продукция являются всенародной государственной собственностью. Благодаря этому при изучении сельскохозяйственного производства совокупность совхозов выступает как качественно однородная статистическая совокупность.

В каждом конкретном статистическом исследовании в соответствии с его познавательными задачами и целями единицы той

¹ Ленин В. И. Полн. собр. соч., т. 30, с. 351.

или иной объективно существующей совокупности могут изучаться с различных позиций, в разных аспектах. Поэтому совокупность, однородная в одном отношении с позиций решения другой познавательной социально-экономической задачи, т. е. при ее изучении в другой связи, может оказаться неоднородной. Так, совокупность промышленных предприятий в СССР качественно однородна в отношении формы собственности. Однако с позиции изучения эффективности производства эта совокупность неоднородна и из нее нужно выделить производственные хозрасчетные объединения, которые обладают существенными преимуществами по сравнению с отдельными предприятиями. Или другой пример. С точки зрения изучения оплаты труда, потребления и т. п. совокупность женщин-работниц, занятых в промышленности СССР, является качественно однородной. С позиции же изучения воспроизводства населения, в частности рождаемости, эту совокупность нельзя признать однородной, так как женщины определенного возраста не могут уже рожать детей.

Наиболее важным в статистике является выделение таких совокупностей, единицы которых однородны в связи со способом производства, например, выделение классов и общественных групп населения, однородных по их отношению к средствам производства. Поэтому при изучении, например, потребления продуктов питания на душу населения советская статистика рассчитывает уровень потребления отдельно в семьях рабочих и в семьях крестьян.

Однако наряду с такими общими статистическими совокупностями, какими являются, например, классы и общественные группы населения, городское и сельское население, колхозы и совхозы и т. п., при решении конкретных познавательных задач выделяются и более частные статистические совокупности: совокупность производственных объединений в промышленности; совокупность женщин плодородного возраста (условно 15—49 лет); совокупность так называемого грамотного населения, а также совокупность населения, способного иметь тот или иной уровень образования (население в возрасте 10 лет и старше) и т. п. В подобных случаях частная статистическая совокупность охватывает либо часть всей совокупности объективно существующих единиц, либо часть единиц общей статистической совокупности, качественно однородной в социально-экономическом отношении.

Для выделения статистических совокупностей (как общих, так и частных) могут быть использованы различные отличительные, характерные особенности общественных явлений и тех элементов, из которых они состоят.

Единицы статистической совокупности обладают многими признаками, т. е. свойствами, отличительными особенностями. Так, отдельные люди отличаются друг от друга по полу, возрасту, семейному положению, образованию и т. п. Каждый признак у отдельных единиц совокупности имеет различные значения или

изменения, которые называются вариантами (например, пол: мужской и женский; возраст: до 1 года, 1,2 и т. д.), а сами признаки — варьирующими.

Варирующие признаки могут быть атрибутивными и количественными. Атрибутивным называется такой признак, варианты которого, характеризующие особенности отдельных единиц, не имеют количественного выражения. Например, при изучении населения атрибутивными признаками людей являются: социальная принадлежность (рабочий, служащий, колхозник), профессия или занятие (слесарь, тракторист, экономист) и т. п.

Количественным называется такой признак, варианты которого выражаются числами (например, возраст, стаж работы, заработная плата).

Вариация изучаемых признаков у отдельных единиц, т. е. их изменение при переходе от одной единицы к другой, — вторая характерная особенность статистической совокупности.

Признаки, которые в пределах изучаемой совокупности не варьируют, т. е. присущи всем ее единицам в одинаковой мере, статистику не интересуют. Так, например, при изучении совокупности колхозов с точки зрения состава их руководящих работников каждый колхоз может быть охарактеризован такими варьирующими от колхоза к колхозу признаками, как число заместителей председателя, число главных специалистов, число агрономов, зоотехников, инженеров и т. п., но было бы бессмысленно ставить вопрос о числе председателей в колхозе, так как в каждом колхозе только один председатель. Аналогично при изучении такой совокупности, как парк тракторов в сельском хозяйстве, каждый трактор можно охарактеризовать такими варьирующими атрибутными признаками, как пахотный данный трактор или нет, колесный он или гусеничный, такими варьирующими количественными признаками, как мощность, год выпуска, степень износа и т. п., но не имеет смысла в перечень изучаемых признаков включать такой, как число тракторов, ибо этот признак не является варьирующим. Характеристика не варьирующих признаков не требует обобщения, так как она ясна непосредственно, на основе изучения отдельной единицы.

Признаки единиц совокупности варьируют не только в пространстве — при переходе от одной единицы к другой, но и во времени. Так, уровень часовой выработки отдельного токаря меняется не только от смены к смене, от одной недели к другой и т. д., но и внутри данной смены. Анализ показывает, что в дневных сменах выработка выше, чем в ночных, что в начале смены и в конце ее выработка ниже, чем в середине смены, и т. п.

Вариация признаков в пространстве и во времени объясняется тем, что на них влияют очень многие обстоятельства и факторы. Так, производительность труда рабочего зависит от его квалификации, качества и состояния оборудования и инструмента, от

уровня организации труда и производства, от материального и морального стимулирования труда и многих других условий.

Урожайность той или иной культуры зависит от качества почвы, качества семян, количества внесенных в почву удобрений, от своевременного проведения других агротехнических мероприятий, от метеорологических условий и т. п.

Таким образом, величина каждого признака у отдельных единиц совокупности обычно складывается под влиянием как других признаков этих единиц, так и под влиянием более общих причин, условий, обстоятельств и факторов, причем все они действуют во взаимосвязи с разной силой и в разных направлениях. В результате этого размеры и количественные соотношения массовых общественных явлений также определяются сложным переплетением взаимодействия различных условий и причин. Одни из этих причин являются внутренними, существенными, вытекающими из природы данного явления, относительно устойчивыми, другие же — внешними, случайными для данного явления или процесса, характеристичными и специфическими только для некоторых единиц совокупности в течение определенного периода времени.

В силу этого то общее, что определяет единство и качественную однородность статистической совокупности и выражает внутренние, существенные связи, т. е. та или иная закономерность, проявляется только в массе всех единиц совокупности, только как *свойство совокупности в целом*.

Такие закономерности обычно выражаются в том, что числовые соотношения размеров массовых общественных явлений обнаруживают *определенный порядок, определенную правильность, последовательность и повторяемость*. Эта правильность может выражаться в относительной устойчивости числовых соотношений размеров явлений в пространстве, т. е. в повторяемости этих соотношений при переходе от одной территории к другой. Известно, например, что между числом мальчиков и девочек среди родившихся наблюдается довольно устойчивое соотношение: на 100 новорожденных приходится в среднем 51—52 мальчика и 48—49 девочек. Такое (или примерно такое) соотношение имеет место довольно длительное время почти во всех странах, если только взять достаточно большую совокупность родившихся. Среди лиц, прибывающих на данную территорию, как правило, мужчин больше, чем женщин. Так, в 1971 и 1972 г. преобладание мужчин среди прибывших имело место по всем союзным республикам СССР.

Определенная закономерность наблюдается также в распределении единиц совокупности по величине некоторых количественных признаков. Так, в населении больше всего людей, рост которых близок к средней величине, причем чем больше рост отклоняется в ту или другую сторону от средней величины, тем меньше численность людей этого роста.

Закономерность может выражаться и в том, что с увеличением размера одного признака размер других признаков изменяется беспорядочно, а с той или иной правильностью. Так, если взять достаточно большую и однородную по почвенно-климатическим и достаточную совокупность колхозов и распределить их на другие условия совокупности количеством удобрений, внесенных на каждый гектар посева, то можно обнаружить, что по мере увеличения количества удобрений от одной группы к другой повышается урожайность. Аналогично по мере увеличения от одной группы рабочих к другой стажа работы повышается их квалификация (тарифный разряд), растет производительность труда, увеличивается средний заработок.

Закономерность может также выражаться в относительной устойчивости количественных соотношений явлений во времени. Например, в СССР в 1969—1971 гг. в среднем на одного жителя приходилось ежегодно 19 посещений платных киносеансов. Гораздо чаще, однако, закономерность проявляется в виде определенной правильности *изменения* с течением времени размеров или числовых соотношений явлений, т. е. в виде их тенденции к росту или к снижению. Так, в нашей стране в течение послевоенного периода из года в год увеличивается доля городского населения в общей его численности, повышается процент лиц с высшим и средним образованием, растут реальные доходы в расчете на душу населения и т. п.

Такого рода закономерности, проявляющиеся только при достаточно большом числе единиц однородной совокупности в условиях конкретного места и времени, т. е. в массовых общественных явлениях, называются *статистическими закономерностями*. Они представляют собой одну из форм проявления всеобщей связи явлений, специфическую форму проявления закономерностей массовых общественных явлений.

В основе статистической закономерности лежит сложное переплетение и взаимодействие внутренних причин, вытекающих из сущности данного явления и общих для всей совокупности однородных единиц, и причин внешних, индивидуальных, обусловленных случайными для каждой отдельной единицы стечением условий и обстоятельств. Именно поэтому статистические закономерности и обстоятельства проявляются только по отношению к массе единиц, а не по отношению к каждой отдельной единице. Здесь проявляется действие закона больших чисел. Сущность этого закона заключается в том, что при достаточно большом числе единиц совокупности и некоторых определенных условиях случайные отклонения от общей меры, свойственные отдельным единицам, взаимно погашаются. В результате такого взаимопогашения отклонений, случайных с точки зрения совокупности в целом, т. е. с точки зрения всей массы, всего массового процесса, проявляется та или иная закономерность явления или процесса. Совокупное действие большого числа случайных факторов приводит к резуль-

тату, почти не зависящему от случая, результату, который и выражает ту или иную статистическую закономерность.

Чем больше число единиц входит в однородную совокупность, тем явственнее выступает закономерность, присущая данному явлению или процессу. Для отдельной же единицы или для небольшого их числа закономерность может не проявиться. Следовательно, закон больших чисел выражает связь между чистой общинной совокупности (числом наблюдений) и степенью, полнотой проявления закономерности, присущей этой совокупности. Так, соотношение между числом родившихся мальчиков и девочек в отдельных семьях может быть самым различным: есть семьи, где рождаются только мальчики, есть семьи, в которых рождаются только девочки, и есть семьи, в которых рождаются и мальчики и девочки в разных соотношениях. Только для достаточно большой совокупности семей обнаруживается довольно устойчивое соотношение, характерное для конкретных условий места и времени, о котором говорилось выше. С точки зрения общей закономерности распределения родившихся по полу рождение мальчика или девочки в отдельных конкретных семьях является случайным, но эти случайности являются той формой, в которой проявляется общая закономерность, необходимо вытекающая из конкретных исторических условий, пока эти условия существенно не изменятся.

Аналогично обстоит дело со смертностью и продолжительностью жизни. Для отдельного конкретного человека невозможно предсказать, когда он умрет и какова, следовательно, будет продолжительность его жизни. Она зависит, с одной стороны, от общих условий жизни в стране, а с другой — от многих местных условий, а также от индивидуальных особенностей жизни и здоровья этого человека. Если же взять достаточно большую группу людей или все население страны, то можно вскрыть определяющие закономерности, которые статистика выражает в специальных таблицах смертности и продолжительности жизни. На основе этих таблиц можно определить, например, какая часть родившихся доживет до определенного возраста, если на протяжении их жизни будут сохраняться условия, характерные для настоящего времени.

Выявление статистических закономерностей, опирающееся на действие закона больших чисел, играет очень большую роль в исследовании общественных явлений. Однако нужно иметь в виду, что закон больших чисел не определяет и не регулирует ни конкретные размеры общественных явлений и процессов, ни их числовое соотношение, ни их изменение во времени. Он лишь обуславливает взаимоположение случайных отклонений и проявления вследствие этого той или иной закономерности. Что же касается содержания этой закономерности, то она определяется сущностью и внутренними законами развития самого явления или процесса. Поэтому выяснение причин той или иной закономерности социально-экономического явления опирается на комплекс-

ное изучение этого явления с помощью ряда наук, предметом которых оно является. Познавание закономерностей развития общественных явлений — это совокупный результат содружества многих наук, такого содружества, которое обеспечивает исследование явлений и процессов не изолированно, а во всем многообразии его взаимосвязей.

4. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ СТАТИСТИКИ

В тесной связи с вопросом о предмете советской статистики находится вопрос о теоретических основах ее, т. е. вопрос о том, с каких научных, теоретических позиций статистика изучает общественные явления и процессы. Советская статистика в своих исследованиях опирается на марксистско-ленинскую теорию. Теоретической основой статистики являются исторический материализм и политическая экономия, которые исследуют и формулируют законы общественного развития. Исходя из принципов и законов этих наук статистика дает конкретную числовую характеристику закономерностям общественных явлений.

Марксистско-ленинская диалектика — методологическая основа статистики

Методологической основой статистики, как и других общественных наук, является марксистско-ленинская диалектика. Это означает, что статистика изучает общественные явления и процессы не изолированно, а в их взаимосвязи, взаимодействии, не в состоянии покоя и неизменности, а в движении, в изменении и развитии. Исходя из законов материалистической диалектики статистика изучает конкретное количественное выражение движущих сил процесса развития общественных явлений, измеряет те границы, за которыми количественные изменения переходят в качественные, дает числовую характеристику направления и тенденций развития.

Замечательные образцы изучения количественных изменений общественных явлений в неразрывной связи с качественными изменениями содержатся в трудах классиков марксизма-ленинизма. Так, в работе «Аграрный вопрос в России к концу XIX века» В. И. Ленин, рассматривая вопрос об аренде земли, выделяет две группы арендаторов. Первая группа — это дворы безлошадные и однолошадные, которые приарендуют одну десятину или даже часть ее; вторая группа — дворы, имеющие четырех и более лошадей и приарендующие 7—16 десятин. «Здесь», — отмечает В. И. Ленин, — количество переходит в качество. Первая аренда есть аренда из нужды, аренда кабатная. «Арендатор», стоящий в подобных условиях, не может не превращаться в орудие эксплуатации посредством отработок, зимней наемки, ссуды денег и т. п. Наоборот, двор, имеющий 12—16 десятин надельной земли, и арендующий сверх того по 7—16 десятин, явно арендует не от нужды, а от богатства, не для «продовольствия для общины», а для того, чтобы «заработать деньги». Мы видим, как во-

очию, — заключает В. И. Ленин, — превращение аренды в капиталистическое фермерство, зарождение предпринимательства в земледелии»¹. Аналогичные примеры перехода количественных изменений в качественные В. И. Ленин приводит также в работе «Развитие капитализма в России».

Классики марксизма-ленинизма неоднократно подчеркивали, что статистика только тогда становится мощным оружием социального познания, когда она опирается на предварительное теоретическое выяснение сущности, законов и форм развития общественных явлений.

Рассматривая вопрос о тенденциях нормы прибыли к понижению, Маркс отмечал, что «Только поняв отношения, действующие при образовании нормы прибыли, статистика приобретает способность предпринять действительный анализ уровня заработной платы в различные эпохи и в различных странах»². В письме к Ф. Энгельсу, рассматривая теорию абсолютной земельной ренты, К. Маркс так охарактеризовал роль статистики и политической экономии в исследовании действия экономических законов. «Единственное, что я должен *теоретически* доказать, это *возможность* абсолютной ренты без нарушения закона стоимости... Что касается *существования* абсолютной земельной ренты, то это такой вопрос, который следовало бы разрешать для каждой страны на *основании статистики*»³.

Критикуя дореволюционную статистику, В. И. Ленин говорил, что нельзя сводить вопрос о развитии крупной машинной индустрии к одной фабрично-заводской статистике, что этот вопрос надо рассматривать шире, как вопрос о тех формах и стадиях, которые проходят развитие капитализма в промышленности страны. Он требовал, чтобы исследователь сначала выяснил сущность этих форм, их отличительные особенности и лишь затем на этой основе иллюстрировал развитие той или иной формы посредством надлежащим образом обработанных статистических данных.

В. И. Ленин предупреждал против превращения статистики в самоцель. «Статистика, — писал Ленин, — должна иллюстрировать установленные всесторонним анализом общественно-экономические отношения...»⁴. На основе такого анализа статистика показывает действие законов политической экономии. Вместе с тем В. И. Ленин подчеркивал, что *политическая экономия должна использовать статистику и статистические данные не только для иллюстрации, но и для проверки и обоснования своих выводов и положений*. Это объясняется тем, что качественные особенности экономических отношений нельзя уяснить без количественного их выражения с помощью статистических данных.

¹ Ленин В. И. Полн. собр. соч., т. 17, с. 88.

² Маркс К. и Энгельс Ф. Соч., т. 25, ч. 1, с. 263.

³ Там же, т. 30, с. 225—226.

⁴ Ленин В. И. Полн. собр. соч., т. 3, с. 506.

4. МЕТОД СТАТИСТИКИ

Опираясь на диалектический метод, а также исходя из характера и основных черт своего предмета, статистика разрабатывает особые, присущие этой науке специальные приемы и способы исследования, составляющие в совокупности статистическую методологию.

Во всяком законченном статистическом исследовании общественных явлений можно различать три этапа или стадии: 1) *статистическое наблюдение*, 2) *сводку материалов статистического наблюдения* и 3) *анализ*. На каждом из этих этапов применяются специфические статистические методы.

На первой стадии статистического исследования происходит собирание первичного статистического материала *методом массового статистического наблюдения*. Значение этого метода заключается в том, что он в соответствии с законом больших чисел позволяет получать достоверные, научно значимые первичные данные об изучаемых общественных явлениях и процессах, дальнейшая обработка которых позволяет вскрыть статистические закономерности.

На второй стадии статистического исследования производится сводка материалов статистического наблюдения в целях выделения однородных статистических совокупностей. Это достигается при помощи ряда методов, важнейшим из которых является *метод статистических группировок*, с помощью которого выделяются качественно однородные социально-экономические типы, характерные по существенным признакам группы и подгруппы.

При сводке материала статистического наблюдения успешно применяется *табличный метод и графики* как способы рационального изложения результатов статистической сводки, облегчающие последующий анализ.

На третьей стадии статистического исследования в процессе анализа и выявления статистических закономерностей и взаимосвязей широко применяется ряд статистических методов, позволяющих получать *обобщающие показатели*, с помощью которых производится измерение, количественная оценка вскрытых при анализе закономерностей.

К таким обобщающим показателям относятся абсолютные и относительные величины, средние величины, показатели вариации, показатели тесноты связи и показатели скорости изменения общественных явлений во времени, индексы.

Широкое применение в статистических исследованиях находят *балансовый метод* как метод обработки и анализа статистических данных, основанный на уравнивании целого и его частей или прихода, расхода и прироста. Различные формы балансов используются для характеристики состава совокупности: соотношений и пропорций между ее отдельными подразделениями и их взаимными

Математические методы

В процессе выявления и анализа статистических закономерностей статистика широко использует также различные математические методы, в частности, такие методы математической статистики, как дисперсионный анализ, регрессионный и корреляционный анализ (см. гл. VIII) и др. Методы математической статистики и теории вероятностей, в частности теоремы закона больших чисел, применяются при изучении общественных явлений не только в процессе обработки и анализа, но и при проведении разного рода выборочных обследований (см. гл. VII).

Подробное изложение существа статистических методов и показателей дается в последующих главах учебника. Здесь же необходимо подчеркнуть, что научная постановка статистического исследования требует тесной взаимной увязки всех трех стадий исследования. Массовое статистическое наблюдение должно быть организовано так, чтобы на основании собранного материала могла быть решена познавательная задача всего исследования.

Сводку материалов статистического наблюдения нужно провести так, чтобы составленные статистические группировки и статистические таблицы могли служить материалом для получения научно обоснованных статистических показателей и последующего их анализа. Наконец, получение статистических показателей и их анализ должны завершаться формулированием научных и практических выводов. Как бы ни был хорош первичный статистический материал, как бы ни была хороша свodka этого материала, но если статистический анализ не дает возможности обосновать конкретные выводы, то вся работа окажется бесполезной.

5. СТАТИСТИКА КАК НАУКА

Определение статистики как науки

Рассмотрение предмета и метода статистики приводит к следующему определению этой науки: статистика — общественная наука, которая изучает размеры и количественные соотношения массовых общественных явлений в конкретных условиях места и времени и дает числовое выражение проявляющимся в них закономерностям. Статистика с помощью чисел показывает степень развития общественных явлений и процессов, скорость их изменения, силу и тесноту их взаимосвязей и взаимозависимостей. Это позволяет утверждать, что статистические показатели (числа) выступают как важнейшее орудие познания разнообразных проявлений общественной жизни.

Поскольку общество является частью единого материального мира, статистика изучает также влияние природных и технических факторов на количественные изменения явлений общественной жизни и обратное влияние развития общественного производства на природные условия жизни общества.

Отрасли статистической науки

Объект статистики — человеческое общество — исключительно сложен. Бурный рост экономики и культуры привел к возникновению отдельных отраслей статистики, необходимых для глубокого и всестороннего изучения общества. Поэтому статистика является многоотраслевой наукой и состоит из отдельных разделов, или отраслей, которые, выступая как самостоятельные ее части, в то же время тесно связаны между собой. В настоящее время получили законченное оформление три отрасли статистики: общая теория статистики, экономическая статистика и отраслевые статистики, т. е. статистики отдельных отраслей народного хозяйства и культуры (промышленности, сельского хозяйства, торговли и т. д.).

Общая теория статистики рассматривает категории статистической науки, а также принципы, правила и методы, которые являются общими для статистического изучения количественной стороны любых массовых общественных явлений.

Экономическая статистика изучает количественную сторону явлений и процессов, происходящих в области экономики, и разрабатывает систему экономических показателей и методы изучения народного хозяйства как единого целого.

Но отдельные отрасли народного хозяйства и культуры (промышленность, строительство, сельское хозяйство, просвещение, здравоохранение и т. п.) имеют свои особенности. Изучение этих особенностей на основе положений общей теории и экономической статистики составляет содержание отраслевых статистик: статистики промышленности, статистики сельского хозяйства, статистики строительства, статистики населения и т. д.

Имея объектом изучения жизнь общества, статистика в той или иной степени связана со всеми общественными науками. Особенно тесной является связь между статистикой и народнохозяйственным планированием. Само составление планов опирается на данные статистики. После составления планов статистика контролирует ход их выполнения, выявляет отстающие участки, резервы перевыполнения плана, обобщает передовой опыт.

Как отрасль общественной науки статистика принципиально, по предмету своего изучения и теоретической основе, отличается от естественных и технических наук, в том числе и от тех, которые носят название статистических (статистическая физика, статистическая механика и др.). В то время как статистика изучает общественные явления и процессы, естественные и технические науки изучают различные явления и процессы органического и неорганического мира.

Статистика отличается также от математики и от математической статистики. Если статистика как общественная наука характеризует размеры и количественные соотношения качественно определенных общественных явлений и процессов в конкретных

условиях места и времени, то математическая статистика, представляющая раздел математики, изучает общие закономерности массовых явлений в абстрактной форме, вне связи со специфической природой изучаемых объектов.

Отличаясь от естественных, технических и математических наук, статистика в то же время не отторгается от них. Всестороннее изучение общества требует учета уровня развития производительных сил и технического прогресса, привлечения математических методов анализа.

Основные задачи статистической науки в СССР определяются прежде всего практическими потребностями государства по управлению и планированию социалистической экономики и культуры. В соответствии с изменением этих потребностей на разных этапах социалистического строительства менялись и задачи статистической науки. В свете решений XXIV съезда КПСС советская статистическая наука призвана обобщать практику и исторический опыт масс в борьбе за построение коммунизма, творчески развивать статистическую теорию, разоблачать апологетический характер буржуазной статистики.

Основной задачей статистической науки является совершенствование и анализ системы показателей, характеризующих строительство коммунистического общества в СССР, т. е. создание материально-технической базы коммунизма, формирование коммунистических общественных отношений и воспитание нового человека. Решение этой задачи связано не только с совершенствованием статистических методов выявления и анализа закономерностей массовых общественных явлений, но и с совершенствованием системы показателей планирования, а также с более тесной увязкой показателей планирования и статистики.

Намеченная партией программа мощного подъема народного хозяйства и повышения благосостояния советского народа ставит перед советской статистикой ряд важных задач. Статистика должна совершенствовать методологию выявления резервов повышения эффективности общественного производства, методологию измерения и анализа эффективности использования трудовых, материальных и финансовых ресурсов, эффективности интенсивных методов ведения хозяйства, а также методологию разработки и анализа системы показателей, характеризующих рост благосостояния и культурного уровня советского народа, постепенное преодоление существенных различий между городом и деревней, между умственным и физическим трудом.

Рост масштабов народного хозяйства и усложнение внутрихозяйственных связей, решение комплексных народнохозяйственных проблем, перспективное долгосрочное планирование, опирающееся на прогнозы роста населения, рост потребностей народного хозяйства и научно-технического прогресса — все это ставит перед статистикой важные задачи в области совершенствования мето-

дологии построения баланса народного хозяйства, совершенствования методов долгосрочного прогнозирования, разработки вопросов создания общегосударственной автоматической системы сбора и обработки информации на основе применения новейшей вычислительной техники.

Критика буржуазных трактовок предмета и содержания статистики как науки Среди буржуазных статистиков широко распространена трактовка статистики как универсального метода изучения массовых явлений природы и общества, метода, являющегося в основном отраслью математики¹. Такое понимание статистики неверно потому, что каждая наука имеет не только предмет изучения, но и свои специфические методы исследования, опирающиеся на диалектический метод и формирующиеся в зависимости от основных характерных черт предмета данной науки. Кроме диалектического метода нет и не может быть универсального метода изучения, одинаково пригодного для всех наук.

Не меньшее распространение среди буржуазных статистиков имеет утверждение, что будто статистика — универсальная наука, являющаяся одним из разделов математики.

Рассматривая статистику как отрасль прикладной математики, буржуазные статистики выдвигают в качестве ее теоретической основы теорию вероятностей².

Конечно, теория вероятностей, если она правильно понимается и применяется, может быть успешно использована при изучении массовых общественных явлений. Но нельзя сводить все изучение, а тем более объяснение общественных явлений к одним только вероятностным схемам. Неумение либо нежелание видеть, что общественные явления подчиняются социальным и экономическим законам, а действие закона больших чисел и других положений теории вероятностей — только форма проявления социально-экономических закономерностей, обрекает буржуазную статистику на научную несостоятельность.

Игнорирование качественного анализа при изучении общественных явлений приводит буржуазную статистику к формально-математическим построениям, являющимся благодатной почвой для различного рода фальсификаций.

Формализм буржуазной статистики нашел наиболее полное воплощение в построениях так называемой англо-американской школы, занимающей господствующее положение в буржуазной статистике. Основателем современного формально-математического направления в статистике является К. Пирсон, идеалистическую

¹ См.: Ю. Л. Дж. Э. и Кендэл М. Дж. Теория статистики (Пер. с англ.). М., Гостатиздат, 1960, с. 17; Дюма Р. Предприятие и статистика (Пер. с франц.). М., Гостатиздат, 1958, с. 9.

² См.: Фишер Р. А. Статистические методы для исследователей (Пер. с англ.). М., Гостатиздат, 1958, с. 11—12; Миллс Ф. Статистические методы (Пер. с англ.). М., Гостатиздат, 1958, с. 7.

сущность философских взглядов которого разоблачил и резко осудил В. И. Ленин.

Ставя вещи с ног на голову, буржуазные статистики рассматривают полученные в результате статистического исследования числовые показатели как эмпирическое воплощение каких-то «идеальных», «истинных» величин, вечных и неизменных по своей природе, независимых от материальных условий общественной жизни и общественного строя. Игнорируя политико-экономический анализ явлений, буржуазная статистика сводит научные изыскания в области общественных наук к схоластически применяемой теории вероятностей.

Глава II

СОВРЕМЕННАЯ ОРГАНИЗАЦИЯ СТАТИСТИКИ

1. УЧЕТ И СТАТИСТИКА В СОЦИАЛИСТИЧЕСКОМ ОБЩЕСТВЕ

Единая система учета и статистики в СССР

В первые же годы после победы Великой Октябрьской социалистической революции одной из важнейших и неотложных задач Советской власти была организация всенародного учета, развивающегося в направлении всестороннего изучения народного хозяйства как единого целого. Эта задача вытекала из характера нового общественного строя, при котором средства производства обобществлены и принадлежат трудящимся.

Под учетом в широком смысле слова принято понимать цифровую характеристику количественной и качественной сторон развития народного хозяйства и других сфер общественной жизни, осуществляемую путем постоянной научно организованной регистрации явлений, происходящих в обществе.

Всестороннее изучение социально-экономических процессов требует применения различных видов учета. Эти виды учета в условиях Советского государства представляют собой единую систему, в которой органически связаны оперативный учет, бухгалтерский учет и статистика.

Оперативный учет — это учет отдельных конкретных фактов и процессов хозяйственной деятельности в целях оперативного руководства предприятиями, организациями и учреждениями. С его помощью осуществляется непосредственная первичная характеристика таких фактов, как явки рабочих на работу, работа отдельных агрегатов, выпуск отдельных видов продукции и т. п. Сведения о таких фактах могут быть получены немедленно вслед за их совершением, и срочность является главным преимуществом оперативного учета по сравнению с другими его видами. Но оперативный учет уступает другим видам учета в вопросах анализа и обобщения. Если некоторое обобщение оперативных сведений и возможно, то анализ явлений только на основе данных оперативного учета неосуществим.

Бухгалтерский учет характеризует кругооборот средств материальных ценностей и другие стороны хозяйственной деятельности предприятий, организаций и учреждений. Основная его цель —

который должен был определять общее направление статистических работ всех министерств и ведомств. Однако для дела такое единство обеспечено не было. В. И. Ленин указывал, что практическая статистика плоха, неслучайно разрозненность между разными ведомствами. «В России... кроме земской статистики, исследовавшей по-европейски небольшие отдельные уголки народного хозяйства, у нас имеется лишь жалкая, перышная, канцелярская-путаная статистика разных ведомств», скорее заслуживающая названия полицейской отписки¹. В. И. Ленин называл канцелярную статистику царской России «горе-статистикой».

Важной организационной формой развития практической статистики в России являлась земская статистика. Она прошла несколько этапов. В первый период (с конца 60-х до середины 90-х годов) земская статистика развивалась относительно самостоятельно, без особого вмешательства правительства. В этот период исследования статистиков земств были направлены на решение широких научных задач. «Первоначально», — писал В. И. Ленин, — земская статистика ограничивалась при переписях данными пообщинности, не собирая данных о составе хозяйств дворян, купцов, однако, заметил различия в наличествующем положении этих дворов и предпринял повороте переписи — это было первым шагом на пути к более глубокому изучению экономического положения крестьян². Начиная со второй половины 90-х годов земские статистики стали проводить обследования, которые позволяли изучать процесс дифференциации крестьянских хозяйств; при разработке итогов переписи на основу стали брать такие признаки, которые характеризовали социальный состав крестьянства и т. д. Данные земской статистики отличались полнотой и достоверностью, так как они собирались квалифицированными специалистами-статистиками, работавшими с научным интересом. По глубине и детальности характеристики экономического положения крестьянства земская статистика не имела себе равной в мире. Проводимые земскими статистиками исследования настолько превосходили работы государственной статистики, что В. И. Ленин считал земскую статистику «замком в пустыне крепостнической темноты, бюрократической рутинки и вселенской тугой канцелярщины»³.

Создание советской государственной статистики и ее совершенствование

Советская государственная статистика начала создаваться в первые же дни после победы Великой Октябрьской революции. Ее начало было положено установлением рабочего контроля. В июне 1918 г. первый съезд статистиков обсудил проект положения о государственной статистике, подписанного В. И. Лениным 25 июля 1918 г. В «Положении» нашли отражение идеи В. И. Ленина о том, что в нашей стране должна быть строгая централизованная, единая система государственных статистических органов, которая обеспечит руководство всеми видами учета в стране. В соответствии с «Положением о государственной статистике» было создано Центральное статистическое управление (ЦСУ). При ЦСУ были созданы Коллегия и Совет по делам статистики, включающие представителей статистических организаций, государственных и общественных учреждений и съездов статистиков. Совет ведал разработкой планов работ и осуществлял методологическое руководство статистическими работами различных ведомств.

Были созданы также статистические органы в губерниях и уездах. Кроме того, для сбора текущих сведений ЦСУ привлекало «добровольных корреспондентов» из местных жителей.

Органы государственной статистики провели большую работу по сбору и обработке статистических данных, необходимых для государственного управления народным хозяйством. Однако в работе государственной статистики были и недостатки. Статистика не всегда полностью отвечала требованиям оперативного руководства и планирования народного хозяйства. Коммунистическая партия провела большую работу по устранению этих недостатков и совершенство-

ванию учета и статистики в стране. Было проведено несколько реорганизаций ЦСУ СССР и его местных органов, проведены мероприятия по механизации статистических работ, организовано производство счетных машин, созданы новые отрасли статистики: материально-технического снабжения, новой техники, природных ресурсов и др.

После XX съезда КПСС права и функции органов государственной статистики были значительно расширены: республиканские статистические управления были преобразованы в ЦСУ союзных республик с предоставлением им широких прав и обязанностей. Была проведена централизация статистики промышленности, строительства и сельского хозяйства. В результате сбора, учета и обработки отчетности по указанным отраслям стали производиться в органах государственной статистики. Тем самым был устранен параллелизм в учете и отчетности.

В конце 50-х годов при всех ЦСУ союзных республик, при крупных краевых и областных статистических управлениях были созданы машинные и вычислительные станции (МЦС), которые в последние годы реорганизованы в вычислительные центры (ВЦ), оборудованные новейшей электронно-вычислительной техникой. Начиная с 1964 г. стали организовываться районные и городские МСС, поэтому в настоящее время имеются почти во всех административных районах страны.

Важной вехой в совершенствовании деятельности органов государственной статистики явилось Положение о ЦСУ СССР, утвержденное Советом Министров СССР в начале 1960 г. Выполнение задач, содержащихся в Положении, обеспечивает значительное улучшение организации учета и статистики, а также повышение роли государственной статистики в улучшении планирования и управления развитием народного хозяйства.

Принципы организации государственной статистики

Принципы организации государственной статистики в СССР определяют ее сущность и закономерности развития социального способа производства.

Важным принципом организации статистики в СССР является неразрывная связь статистических организаций с органами государственного управления и планирования. Этот принцип вытекает из указаний В. И. Ленина, который подчеркивал, что «оно (Центральное статистическое управление. — Авт.) должно быть не «академическим» и не «независимым» органом, а органом социалистического строительства, проверки, контроля, учета того, что надо социалистическому государству знать теперь, сейчас, в первую голову»¹. В соответствии с этим принципом в нашей стране была создана централизованная система учета и статистики, которая является важнейшим рычагом государственного управления и планового руководства народным хозяйством, обеспечивающим строгий контроль за выполнением государственных планов и соблюдением общегосударственных интересов. Статистические органы в нашей стране являются специализированными управлениями при Совете Министров СССР и советах министров союзных республик.

Другим принципом организации советской статистики является строгая централизация руководства делом учета и отчетности в стране. Обобществление средств производства, социалистический

¹ Ленин В. И. Полн. собр. соч., т. 12, с. 354.

² Там же, т. 1, с. 9.

³ Там же, т. 27, с. 182.

Ленин В. И. Полн. собр. соч., т. 45, с. 156.

характер Советского государства позволили создать систему государственной статистики, в которую, наряду с общегосударственной статистикой (ЦСУ и его местные органы), входят статистические органы министерств и ведомств (ведомственная статистика). Центральное статистическое управление осуществляет централизованное руководство учетом и статистикой в нашей стране. Все звенья разветвленной системы статистических органов строго централизованы и работают по единому плану.

Современная организация советской статистики.

ЦСУ СССР, его права и обязанности

Вступление нашей страны в период развитого социалистического общества потребовало совершенствования системы и методов управления и планирования, а также дальнейшего совершенствования советской статистики и ее организации. Важную роль в этом должно сыграть новое Положение о Центральном статистическом управлении при Совете Министров СССР (ЦСУ СССР), которое утверждено Советским правительством в апреле 1973 г.

В этом Положении сохранены ленинские принципы организации государственной статистики при социализме и в соответствии с решениями XXIV съезда КПСС четко определена роль статистики в осуществлении мероприятий по дальнейшему развитию хозяйственной реформы в стране и совершенствованию управления и планирования в народном хозяйстве.

ЦСУ СССР является союзно-республиканским органом, осуществляющим централизованное руководство делом учета и статистики в СССР.

Главной задачей ЦСУ СССР является сбор, разработка, анализ и своевременное представление Совету Министров СССР, Госплану, министерствам и комитетам Совета Министров СССР достоверных, научно обоснованных статистических данных, которые характеризуют ход выполнения государственных планов, эффективность общественного производства, научно-технического прогресса и другие стороны экономической и культурной жизни страны. С помощью статистических данных ЦСУ СССР должно показывать наличие и использование производственных мощностей, природных, трудовых и материальных ресурсов, выявлять имеющиеся в народном хозяйстве резервы для перевыполнения государственных планов, а также изучать происходящие в стране социально-экономические и научно-технические процессы.

Для решения этих задач ЦСУ СССР должно совершенствовать научную методологию и организацию учета и статистики и систему статистических показателей, расширять сферу применения балансового метода, а также экономико-математических методов и электронно-вычислительной техники.

Исходя из указанных задач Совет Министров СССР возложил на ЦСУ СССР следующие обязанности:

осуществление централизованного руководства делом учета и статистики и организацию учета и контроля за ходом выпол-

нения государственных планов развития народного хозяйства; организацию и ведение статистики отраслей народного хозяйства и культуры, а также других статистических работ, в том числе различных переписей и обследований;

обеспечение создания автоматизированной системы государственной статистики (АСГС) и широкого внедрения во все отрасли народного хозяйства электронно-вычислительной техники для сбора и обработки учетно-отчетной документации;

утверждение общесоюзного минимума показателей и форм статистической отчетности для предприятий и организаций и учреждений, а также разработку и издание инструкций к заполнению форм статистической отчетности;

общее руководство вопросами первичного учета и проверки правильности его организации на предприятиях, в учреждениях, организациях, министерствах и ведомствах;

контроль за состоянием учета и отчетности и достоверностью отчетных данных во всех отраслях народного хозяйства, дальнейшее сокращение отчетности, сроков и удешевление стоимости ее разработки, а также недопущение отчетности, не утвержденной в установленном порядке;

внедрение и эффективное использование экономико-математических методов и вычислительной техники при разработке статистических данных, осуществление методологического руководства развитием механизации и автоматизации учетно-вычислительных работ в народном хозяйстве;

публикацию в печати сообщений о выполнении государственных планов развития народного хозяйства СССР, издание статистических сборников, журнала «Вестник статистики» и документов по учету и статистике;

руководство делом подготовки и переподготовки массовых кадров по статистической специальности и по механизации учетно-статистических работ;

систематизацию статистических данных по экономике и культуре зарубежных стран; представительство в международных статистических организациях, на съездах и конференциях;

осуществление взаимного обмена опытом работы со статистическими организациями социалистических стран.

ЦСУ СССР предоставлено право отменять отчетность, не утвержденную в установленном порядке. Указания ЦСУ СССР об отмене такой отчетности являются обязательными для министерств и ведомств, исполкомов Советов депутатов трудящихся, а также для всех предприятий, строек, колхозов, совхозов, учреждений и организаций.

ЦСУ СССР и органы государственной статистики имеют право требовать от предприятий, строек, колхозов, совхозов, учреждений, министерств и ведомств статистическую и бухгалтерскую отчетность и материалы по всем учетно-статистическим работам в любой стадии их разработки, а также дополнительные данные и изменения к представляемой отчетности.

Структура ЦСУ СССР

Во главе Центрального статистического управления при Совете Министров СССР стоит начальник, назначаемый Верховным Советом СССР.

В ЦСУ СССР образована коллегия из состава начальников ЦСУ СССР (председатель), заместителей и руководителей работников ЦСУ. Состав коллегии утверждается Советом Министров СССР. Коллегия рассматривает важнейшие вопросы организации и методологии статистических работ, итоги основных статистических работ, вопросы централизации и механизмы учета и отчетности, проверки исследования, подбора и расстановки кадров, заслушивает отчеты руководителей работников управлений и отделов ЦСУ СССР и местных статистических органов.

Для разработки и обсуждения методологических вопросов статистики, программ, инструкций по важнейшим статистическим работам при ЦСУ СССР работает научно-методологический совет. Члены совета утверждаются коллегией ЦСУ СССР из числа научных и практических работников в области статистики.

Аппарат ЦСУ СССР состоит из ряда управлений и большого числа отделов, организованных преимущественно по отраслям народного хозяйства. В ЦСУ СССР функционируют управления статистики промышленности, сельского хозяйства; отделы статистики торговли, транспорта и т. д. Внутри управлений также имеются отделы. Например, в управлении статистики сельского хозяйства функционируют отделы агро-зоо-экономической, статистики земледелия и урожайности, статистики животноводства, статистики труда и себестоимости, статистики совхозов, статистики заготовок и др. Наряду с отделами, руководящими статистикой отдельных отраслей народного хозяйства, имеются сводные отделы, которые или охватывают народное хозяйство в целом, или занимаются отдельными вопросами, касающимися всего народного хозяйства. Например, отделы баланса народного хозяйства, статистики труда, информации, сводной статистики и статистической методологии и др.

В ведении ЦСУ СССР находятся и действуют на основе хозяйственного расчета: Главное управление вычислительных работ (Главмехсчет), Всесоюзное производственно-техническое объединение по техническому обслуживанию и ремонту вычислительной техники (Совзестехника ЦСУ СССР), Всесоюзный государственный проектно-технологический институт по механизации учета и вычислительных работ ЦСУ СССР и Главное управление подготовки и повышения квалификации работников учета (ГУПК ЦСУ СССР).

При ЦСУ СССР работает Научно-исследовательский институт по проектированию вычислительных центров и систем экономической информации (НИИ ЦСУ СССР). Институт разрабатывает вопросы создания автоматизированной системы государственной статистики (АСГС) и единой системы вычислительных центров, оснащенных новейшими средствами хранения, обработки и передачи экономической информации. В 1974 г. создан Научно-технический совет по проблемам проектирования и создания АСГС.

Для централизованной разработки статистической информации в системе органов государственной статистики имеются Главный вычислительный центр, республиканские вычислительные центры и машинно-вычислительные станции (МСС).

В 1971 г. началось объединение районных (городских) институтов государственной статистики с районными (городскими) МСС системами ЦСУ СССР, которое должно быть завершено в ближайшие годы. В результате такого объединения организуется единый орган — районная (городская) информационно-вычислительная станция государственной статистики (РИБС, ГИВС). Образование РИБС (ГИВС) является новым, важным шагом на пути дальнейшего совершенствования государственной статистики.

Органы государственной статистики в республиках, краях, областях, районах и городах

В Положении о ЦСУ СССР указывается, что делом учета и статистики ведают: в союзных республиках — *центральные статистические управления при советах министров союзных республик*, в автономных республиках, краях, областях, округах,

в Москве, Ленинграде и столицах союзных республик — *статистические управления*.

В районах и городах учетом и статистикой ведают *районные (городские) информационно-вычислительные станции государственной статистики*.

Центральные статистические управления при советах министров союзных республик являются союзно-республиканскими органами, которые осуществляют руководство делом учета и статистики на территории союзных республик и подчиняются как советам министров союзных республик, так и ЦСУ СССР. ЦСУ союзных республик выполняют статистические работы, возлагаемые на них постановлениями и распоряжениями Совета Министров СССР, централизованным общесоюзным планом статистических работ, постановлениями и распоряжениями советов министров союзных республик и в порядке отдельных заданий ЦСУ СССР.

Центральные статистические управления при советах министров союзных республик, статистические управления автономных республик, краев и областей выполняют на своей территории те же функции (в отношении учета и статистики), что и ЦСУ СССР в масштабах страны, т. е. содействуют составлению соответствующих территориальных планов, осуществляют контроль за их выполнением, проводят единовременные статистические работы, собирают и разрабатывают отчетность и т. д.

Первичным звеном централизованной системы ЦСУ СССР являются районные (городские) информационно-вычислительные станции государственной статистики. На них возлагается:

выполнение статистических работ, предусмотренных постановлениями и распоряжениями Правительства СССР и союзной республики, централизованным общесоюзным планом статистических работ и отдельными заданиями ЦСУ СССР, ЦСУ союзной республики и статистических управлений автономных республик, краев и областей;

разработка учетно-отчетной документации предприятий, колхозов и организаций;

получение и разработка отчетности предприятий, строек, колхозов, совхозов, учреждений и организаций, а также данных переписей, единовременных учетов и статистических обследований;

выпуск статистических материалов и представление их местным руководящим и финансовым органам, а также плановым комиссиям; экономический анализ статистической информации и внесение на рассмотрение местных руководящих органов вопросов, связанных с развитием хозяйства и культуры;

систематическая проверка учета и отчетности на предприятиях, стройках, в колхозах, совхозах, учреждениях и организациях и оказание им помощи в учетно-статистической работе; обеспечение своевременной и доброкачественной отчетности;

выявление и отмена отчетности, не утвержденной в установленном порядке;

систематизация статистических данных о развитии хозяйства и культуры за продолжительный период.

На основании годового плана статистических работ, составленного ЦСУ СССР, статистические управления республик, краев и областей поквартально разрабатывают план работы РИВС (ИВРС).

Основные публикации
органов государ-
ственной статистики

Основная публикация органов государственной статистики

ЦСУ СССР регулярно публикует статистическую информацию в виде сообщений об итогах выполнения квартальных, полугодных и годовых государственных планов, развития народного хозяйства СССР, статистических ежегодников «Народное хозяйство СССР в . . . году» и кратких ежегодных «Справок в цифрах в . . . году».

Сборники статистических сборников ЦСУ СССР содержат данные не только за

«Народное хозяйство СССР» содержит данные не только за

[illegible]

Известны также статистические сборники по отдельным отраслям народного хозяйства, по отдельным вопросам и итогам крупных статистических работ (Госкомстатиздательство СССР, «Сельское хозяйство СССР», «Итоги Всесоюзной переписи населения 1970 года» и т. п.).

ЦСУ СССР издает также ежемесячный журнал «Вестник статистики», в котором печатаются статьи по актуальным вопросам теории и практики советской статистики, даются консультации по вопросам учета и статистики, а также проводятся публикации статистических материалов.

Местные органы ЦСУ публикуют статистические данные по соответствующим территориям.

Организация ведомственной статистики

На промышленных предприятиях, стройках, колхозах, совхозах, организациях, государственных комитетах, министерствах и ведомствах имеются отделы, секторы, или отдельные работники, которые ведут статистическую работу в пределах этих предприятий, учреждений и организаций. Такие статистические органы называются органами ведомственной статистики. Их работники собирают и обрабатывают статистические данные, руководят деятельностью предприятий, организаций и властей, а также для составления внутриведомственных планов и контроля за их выполнением.

Статистические органы (или отдельные работники) предприятий, организаций и учреждений находятся в двойном подчинении: в административном отношении они подчиняются ведомству, а в методологическом — органам ЦСУ СССР.

В первую очередь необходимо обеспечить единство учета и отчетности, а также сами формы ведомственной отчетности. В первую очередь это касается бухгалтерского учета. В настоящее время в СССР нет единой формы бухгалтерского учета, что приводит к тому, что бухгалтерские отчеты предприятий, организаций и учреждений не могут быть сопоставлены между собой. Это приводит к тому, что бухгалтерские отчеты предприятий, организаций и учреждений не могут быть сопоставлены между собой. Это приводит к тому, что бухгалтерские отчеты предприятий, организаций и учреждений не могут быть сопоставлены между собой.

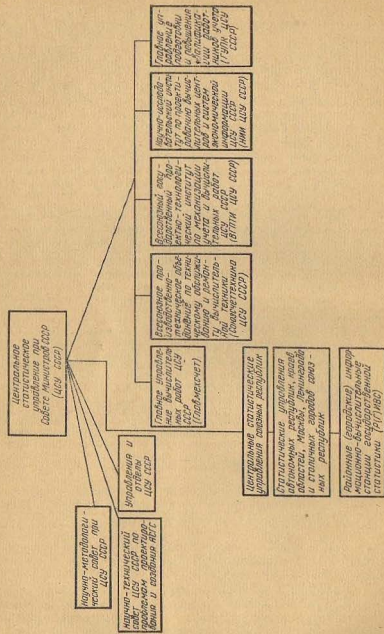


Рис. 2.1. Структура централизованной системы органов государственной статистики в СССР

3. ОРГАНИЗАЦИЯ СТАТИСТИКИ В ЗАРУБЕЖНЫХ СОЦИАЛИСТИЧЕСКИХ СТРАНАХ

В зарубежных социалистических странах с первых дней победы революции большое внимание уделяется вопросам создания и развития марксистско-ленинской статистики. Созданная в этих странах общегосударственная статистика основана на ленинских принципах и развивается с учетом многолетней практики советской статистики.

Все зарубежные социалистические государства имеют в системе органов управления и планирования *центральные статистические организации*, которые по своим функциям и структуре имеют много общего с ЦСУ СССР. Например, в Польской Народной Республике руководит учетом и статистикой Центральное статистическое управление при Совете Министров, в Чехословацкой Социалистической Республике — Федеральное статистическое управление и т. д.

Центральные статистические органы социалистических стран имеют несколько различную структуру, но им присуще много общего. В большинстве стран центральный статистический орган имеет главные отделы по отраслям народного хозяйства и отделы: отчетности, научной публикации, бюджетов и др.

В каждой зарубежной социалистической стране имеются и *местные статистические органы* (управления, инспектуры или бюро, отделы местных органов власти), построенные в соответствии с административно-территориальным делением страны.

Как и в СССР, в других социалистических странах наряду с общественной статистикой имеется *ведомственная статистика*. В министерствах и ведомствах функционируют либо самостоятельные статистические отделы, либо подотделы и секции плановых отделов. На предприятиях имеются статистические секции при плановых отделах или отдельные работники статистики. Статистические органы большинства социалистических стран осуществляют методологическое руководство ведомственной статистикой. Они рассматривают и утверждают программы и формы ведомственной отчетности, контролируют правильность составления и достоверность отчетных данных.

Основные задачи, важнейшие работы и публикации органов статистики

Статистику стран мировой социалистической системы роднит единство целей и задач. Во всех социалистических странах органы статистики обслуживают интересы народа и потребности государственного планового руководства. Основными задачами органов государственной статистики в зарубежных социалистических странах являются: современное и точное отображение в статистических данных всех сторон социалистического строительства; разработка системы научно обоснованных статистических показателей, характеризующих развитие народного хозяйства;

удовлетворение потребностей планирующих органов в необходимых данных для разработки народнохозяйственных планов; осуществление контроля за выполнением плана и анализ причин отклонений от директивных показателей; обеспечение сбора, разработки и анализа достоверных статистических данных и своевременное представление их правительству и различным государственным учреждениям; руководство механизацией учета и отчетности; проведение общегосударственных переписей и других единовременных статистических работ, имеющих общегосударственное значение; публикация различных статистических материалов; участие в разработке теоретических вопросов экономики, планирования и статистики.

Благодаря использованию многолетнего опыта советской статистики за сравнительно короткий срок зарубежные социалистические страны добились больших успехов в развитии марксистско-ленинской статистики и проведении статистических работ. Во всех странах создана текущая статистика, показывающая ход выполнения народнохозяйственных планов. Систематически проводятся учеты и переписи населения, промышленности и других отраслей народного хозяйства и культуры. В большинстве стран проводятся выборочные обследования бюджетов общественных групп населения, составляются балансы народного хозяйства.

Важная работа ведется по механизации учетно-статистических работ. В настоящее время во всех социалистических странах для обработки статистической информации применяется новейшая электронно-вычислительная техника. Особого внимания заслуживает опыт механизации учетно-статистических работ, осуществляемой в ГДР. Здесь во всех сферах народного хозяйства разработаны одинаковые формы документов с учетом различной степени механизации обработки данных, создана центральная служба первичной отчетности, которая отвечает за разработку стандартов для форм отчетности. Сроки механизации получения и разработки статистической информации с помощью ЭВМ предусматриваются в перспективном плане развития учета и статистики во всех отраслях народного хозяйства на период до 1975 г., а по отдельным вопросам — до 1980 г.

Во всех социалистических странах получили широкое развитие статистические публикации и другие статистические издания. Центральные и местные статистические органы всех стран регулярно публикуют сообщения о выполнении народнохозяйственных планов. Большое место в издательской деятельности статистических органов занимают результаты статистических исследований и вопросы методологии статистических наблюдений. Так, в ГДР издаются серии «Статистические исследования и работы», «Методологические тетради», «Массовые статистические обследования».

Одной из главных причин больших успехов зарубежных социалистических стран в развитии статистики является интенсивная подготовка статистических кадров. Центрами подготовки статистиков являются специализированные высшие школы. Так, в

СССР подготовка статистических кадров сосредоточена в высших экономических школах — Пражской и Братиславской, в ГНР основным учебным заведением, готовящим статистиков, является институт планирования и статистики в Варшаве и т. п. В этих учебных заведениях будущие статистики получают широкое экономическое образование.

Сотрудничество между странами мировой социалистической системы в области статистики

Успешному развитию статистики в социалистических странах способствует систематический обмен опытом работы между статистическими органами этих стран. Статистики зарубежных социалистических стран часто приезжают в СССР для изучения богатейшего опыта, накопленного советскими статистиками, а советские статистики, выезжая в другие социалистические страны, изучают опыт работы статистических органов этих стран. Между государствами мировой системы социализма налажен систематический обмен статистическими изданиями, публикациями и т. д.

Развитие экономического сотрудничества между странами мировой социалистической системы потребовало значительного усиления сотрудничества и в области статистики. Важной вехой на пути укрепления сотрудничества социалистических стран в области статистики явилось создание в 1962 г. Постоянной комиссии по статистике, являющейся органом СЭВ. Основные задачи комиссии сводятся к разработке рекомендаций по дальнейшему совершенствованию организации и методологии статистики, повышению уровня статистической работы и экономического анализа, а также разработке сравнимых статистических данных, необходимых для работы органов СЭВ.

В 1971 г. XXV сессия СЭВ приняла Комплексную программу дальнейшего углубления и совершенствования сотрудничества и развития социалистической экономической интеграции стран — членов СЭВ, рассчитанной на 15–20 лет. В этой программе подчеркнуто, что социалистические страны «... в области взаимной статистической информации будут согласовывать основные методологические принципы разработки статистических показателей, а также их сопоставимость».

Заседания комиссии проводятся два раза в год. Рекомендации, принимаемые на этих заседаниях обязательны для всех стран — членов СЭВ. Такое положение способствует унификации статистического учета. Подготовкой рекомендаций для комиссии занимаются рабочие группы специалистов-статистиков, созданные по различным отраслям статистики.

4. ОРГАНИЗАЦИЯ СТАТИСТИКИ В КАПИТАЛИСТИЧЕСКИХ СТРАНАХ

Развитие капитализма, его превращение в монополистический, загнивающий и умирающий капитализм непрерывно усиливало централизацию исполнительной власти, а это в свою очередь ве-

ло и возрастанию значения централизованных статистических органов. Однако чем больше в капиталистическом обществе усугубляются противоречия и обостряется классовая борьба, тем больше статистика становится орудием буржуазной апологетики, орудием фальсификации.

Государственная статистика капиталистических стран со стороны организационной, со стороны полноты и качества полученных ею материалов и, наконец, с точки зрения их обработки, анализа и его результатов стоит неизмеримо ниже государственной статистики СССР и других социалистических стран.

В капиталистических государствах нет единообразной и стройной системы организации статистических учреждений. Органы статистики там разрознены. Статистической работой в капиталистических странах занимаются как частные организации, так и государственные учреждения, которые собирают сведения, необходимые капиталистическому государству для его административной деятельности. Частные же статистические учреждения (частные исследовательские институты и т. д.) собирают сведения, необходимые для обеспечения интересов той или иной группы лиц.

Деятельность всех учреждений и организаций, занимающихся сбором и публикацией статистических данных, обычно не координируется. Каждое статистическое учреждение собирает сведения по такому кругу объектов, о таких показателях и такими методами, которые оно считает наиболее приемлемыми. Это приводит к тому, что по ряду важнейших показателей сведения отсутствуют, а из сведений, собранных по некоторым показателям, из-за их несопоставимости невозможно получить данные не только по стране в целом, но и по отдельным отраслям народного хозяйства.

Во многих капиталистических странах существуют центральные правительственные статистические бюро. Но «центральными» они во многих случаях только называются. Фактически эти бюро играют роль только совещательных и наблюдательных органов, ибо всей статистической работы в стране не объединяют, а занимаются определенными статистическими операциями, например переписями, а также подготовкой публикаций на основе материала, собираемого различными учреждениями. Такое положение «центрального» статистического органа приводит не только к дублированию одних сведений и недоучету других, но и к дублированию самой работы по сбору сведений.

Государственная статистика капиталистических стран страдает разрозненностью, раздробленностью, отсутствием планомерной организации. Особенно отрицательно на статистической работе государственных учреждений сказывается принцип добровольности в предоставлении сведений предпринимателями.

В странах капитала из-за отсутствия единства в самой системе хозяйства, раздираемой антагонистическими противоречиями и яростной капиталистической конкуренцией, нет и не может быть

единой системы учета и статистики. При капитализме учет совершенно оторван от статистики. Капиталисты широко развивают учет внутри предприятий. Этот учет обслуживает интересы отдельных капиталистов: помогает управлять предприятием, эксплуатировать рабочих, подсчитывать прибыли. Однако подавляющая часть данного учета не обобщается статистикой. Капиталисты строго охраняют «коммерческую тайну» и поэтому принимают меры к тому, чтобы утаить сведения о работе своего предприятия.

Для того чтобы скрыть истинный размер получаемых прибылей, капиталисты вообще отказываются давать ряд сведений. В связи с этим буржуазные статистики часто прибегают к различным косвенным расчетам, дающим широкий простор для тенденциозного изображения капиталистического хозяйства.

Необходимость сохранения коммерческой тайны вносит ограничения и в публикацию статистических данных. Так, в США итоговые статистические данные, если они относятся к одной или двум фирмам, не могут быть опубликованы. И только в том случае, если статистические материалы охватывают три фирмы и более, их можно опубликовать при условии, что на долю одной или двух из них приходится менее 90% общего итога. В опубликованных статистических материалах часто содержится прямая фальсификация данных. При этом чем больше политическое значение и остроту имеет исследуемый вопрос, тем больше извращений вносит в его изучение буржуазная статистика. Более всего извращаются данные о безработице, о заработной плате, о стоимости жизни, о прибылях капиталистов и т. д.

В организации статистики капиталистических стран имеется много общего, но есть и различия. В качестве примера рассмотрим организацию статистики в США. В этой наиболее высокоразвитой стране статистическая работа ведется многочисленными учреждениями и лицами. Наряду с учреждениями федерального правительства, статистикой занимаются правительственные учреждения отдельных штатов, а также общественные организации.

Основными государственными органами, которые занимаются сбором статистических данных, являются: Бюро переписи, Бюро статистики труда, Служба сбора сельскохозяйственных продуктов, Национальное управление статистики естественного движения населения. Однако между статистическими учреждениями нет четкого разграничения сферы работы и задач, отсутствует единая методология. В результате по одному и тому же вопросу каждый статистический орган дает различные показатели. Например, при учете численности работающих в сельском хозяйстве министерство сельского хозяйства относит к занятым всех работающих подростков, а Бюро переписи — только лиц старше 14 лет и т. п. В результате расхождения в данных о численности работающих достигают 25%.

Координацию работ в области статистики в США должно осуществлять Управление статистических стандартов. Однако в условиях капиталистической действительности оно не может выполнять возложенные на него задачи и поэтому его функции в основном остаются чисто консультативными. Отдельные штаты в отношении выполнения статистических работ действуют самостоятельно, не подчиняясь от федерального правительства.

Основными статистическими работами в США являются так называемые переписи (переписи). До недавнего времени американские переписи охватывали население и следующие отрасли хозяйства: сельское хозяйство, торговлю и услуги. Поэтому они представляли собой очень обширные статистические операции, в которых принимало участие огромное число лиц. В последнее время переписи населения стали производиться отдельно.

Основным методом производства переписей промышленности, сельского хозяйства и других отраслей в США является рассылка анкет предпринимателям, фермерам. В таких условиях точность даже простейших показателей зачастую оказывается невысокой. Кроме того, капиталисты — руководители предприятий не стремятся ни к точности данных, ни к своевременности представления анкет. Например, в 1957 г. при обследовании обрабатывающей промышленности американцы прислали сведения лишь 30% обследуемых предприятий. Все это снижает достоверность статистики США все чаще прибегают к выборочному методу.

В США по сравнению с другими капиталистическими странами публикация статистических данных производится наиболее широко. Государственными органами ежегодно публикуется свыше 200 официальных статистических изданий. Наиболее важным из них является ежегодный министерства торговли «The Statistical Abstract of the United States», содержащий данные по экономике страны; ежеквартальный журнал министерства торговли «Survey of Current Business», охватывающий основные сферы экономики США; ежемесячный журнал федерального резервного управления «Federal Reserve Bulletin», в котором публикуются экономические статьи и статистические данные, характеризующие экономику США, и др.

МЕЖДУНАРОДНЫЕ СТАТИСТИЧЕСКИЕ ОРГАНИЗАЦИИ

Развитие экономических и культурных связей между государствами и образование мирового хозяйства привели к созданию международных статистических организаций. Первыми такими организациями были международные статистические конгрессы.

Было проведено десять заседаний конгрессов: первый — в Брюсселе в 1853 г., а последний — в Буапенте в 1876 г. Являясь полуофициальной международной организацией, конгрессы оказали значительное влияние на развитие статистической науки и совершенствование практики обработки и публикации данных международной статистики. В работе конгрессов активное участие принимали русские статистики П. П. Семенов-Тянь-Шанский, Ю. Э. Ясон и др.

В 1885 г. был основан *Международный статистический институт* (МСИ). С момента организации состоялось 38 сессий института. Основная задача МСИ определена его уставом, где говорится, что «международный статистический институт является ставившей целью, имеющей целью развитие и усовершенствование статистических методов и их применение в различных странах мира». Работы сессий института печатаются в «Бюллетене института», а другие материалы — в «Вестнике Международного статистического института». До первой мировой войны МСИ был главной международной статистической организацией, занимавшейся не только разработкой вопросов методологии международной статистики, но и обеспечением сопоставимости и публикации статистических данных.

Значительные изменения в функциях МСИ были вызваны созданием статистических органов сначала при Лиге Наций, а затем при Организации Объединенных Наций. Ныне институт занимается разработкой и совершенствованием методов статистики, главным образом математической, отчасти демографической и экономической. В состав МСИ входят советские ученые академики А. И. Козмогоров и Ю. В. Линник, действительный член АН УССР В. В. Гнеденко и член-корреспондент АН СССР Т. В. Рыбушкин.

В настоящее время основным методологическим органом международной статистики является Статистическая комиссия Организации Объединенных Наций. Она создана в 1946 г. и подчиняется Экономическому и Социальному Совету ООН. Статистическая комиссия занимается разработкой вопросов статистической методологии и сопоставимости данных, выработкой рекомендаций для Статистического бюро Секретариата ООН, координацией статистической работы специализированных органов ООН и консультативным их по вопросам собирания, обработки и анализа статистических данных. Для обсуждения вопросов международной статистики комиссия собирается на сессии.

Исполнительным статистическим органом ООН является Статистическое бюро Секретариата ООН. Бюро собирает, обрабатывает и публикует данные по международной статистике, получаемые от статистических органов государств—членов ООН и специализированных органов ООН. Оно имеет ряд периодических изданий: а) ежемесячный статистический бюллетень, данные которого поступают более чем из 140 стран мира; б) статистический ежегодник, данные для которого представляют свыше 150 стран; в) демографический ежегодник; г) ежегодник по внешней торговле; д) ежеквартальный статистический журнал и др.

Вопросами статистики занимаются также региональные экономические комиссии ООН для Европы, Азии и Дальнего Востока, Латинской Америки. Определенное значение в организации международной статистики имеют постоянно действующие конференции европейских, азиатских и африканских статистиков. Их задачами являются разработка методологических проблем статистики, содействие улучшению качества и обеспечению сопоставимости статистических данных, а также налаживание сотрудничества международных организаций в области статистики. Сессии конференций проводятся один раз в год.

Статистическая работа в ООН ведется и международными специальными органами: Международной организацией труда, Продовольственной и сельскохозяйственной организацией ООН, ЮНЕСКО, Международной организацией здравоохранения и др. Эти организации занимаются сбором, обработкой и публикацией статистических данных по соответствующим отраслям.

Большинство международных статистических организаций в той или иной степени способствуют развитию статистической науки, обеспечению сопоставимости статистических данных и совершенствованию статистической практики. В ряде рекомендаций, разработанных этими организациями, частично отражены пожелания статистиков социалистических стран. Однако эти рекомендации не свободны от неправильной методологии буржуазной статистики.

В настоящее время есть и такие международные статистические организации, которые своей задачей ставят не развитие международной статистики, а апологетику капитализма. Подобной организацией является, например, Международная ассоциация по изучению национального дохода и национального богатства, финансируемая фондом Рокфеллера.

Глава III

СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ

СТАТИСТИЧЕСКОЕ НАБЛЮДЕНИЕ И ЕГО ЗАДАЧИ

Наблюдение — первый этап статистического исследования

Для того чтобы изучить количественную сторону явлений и процессов общественной жизни, прежде всего нужно собрать о них необходимые данные. Поэтому первым этапом каждого статистического исследования является статистическое наблюдение. Оно играет особую роль, так как ошибки, допущенные в процессе собирания данных, очень трудно, а иногда и невозможно исправить на дальнейших этапах работы. Поэтому при организации и проведении статистического наблюдения необходимо очень строго соблюдать правила и требования статистической науки.

Статистическое наблюдение представляет собой планомерное, научно организованное собирание массовых сведений о социально-экономических явлениях и процессах.

Статистическое наблюдение — это большая, сложная и трудоемкая работа, обычно выполняемая силами многих работников. Оно складывается из подготовительных работ, непосредственного собирания данных и их контроля. Важнейшим элементом подготовки наблюдения является разработка плана его проведения, который входит в план всего статистического исследования. При составлении этого плана всегда необходимо иметь в виду конечную цель исследования, т. е. решение поставленных перед ним задач. Поэтому план наблюдения должен разрабатываться так, чтобы было обеспечено получение всех данных, которые требуются на последующих этапах работы.

Одной из важнейших задач статистического наблюдения в СССР и других странах мировой социалистической системы является получение данных, необходимых для

составления народнохозяйственных планов и контроля за их выполнением. Отсюда вытекает неразрывная связь показателей плана и показателей статистики и необходимости их сопоставления. Однако круг показателей статистики шире круга плановых показателей, так как в число показателей статистики включаются данные, которые неуместны в плане. Например, в промышленной

статистике собираются сведения о прогулах, о текучести рабочей силы и т. п., которые, разумеется, не планируются. Кроме того, статистика использует ряд показателей, необходимых для выяснения причин невыполнения плана или для выявления неиспользованных резервов и распространения передового опыта. Например, статистика собирает сведения о численности рабочих, не выполняющих, выполняющих и перевыполняющих нормы выработки, о причинах простоев тракторов и т. п.

Статистическое наблюдение должно также обеспечивать получение данных, необходимых для различных социально-экономических исследований, выявления и изучения закономерностей и тенденций в развитии общественных явлений.

Собираемые сведения должны быть достоверными. Характер общественного устройства стран мировой социалистической системы определяет преимущество статистического наблюдения в странах социализма перед наблюдением, проводимым в буржуазных странах. В нашей стране малейшее искажение данных статистического наблюдения считается уголовно наказуемым деянием. Буржуазные же статистики в угоду империализму при проведении статистического наблюдения принимают все зависящее от них для того, чтобы затуманить противоречия, присущие капиталистическому способу производства.

2. ФОРМЫ, ВИДЫ И СПОСОБЫ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ

Статистическое наблюдение может быть организовано по-разному. В статистике различают две организационные формы наблюдения: отчетность и специально организованное статистическое наблюдение.

Отчетность предприятий, учреждений и организаций является основной формой собирания статистических сведений в СССР и других странах мировой социалистической системы. Отчетность представляет собой такую организационную форму статистического наблюдения, при которой в установленные сроки и по определенным адресам представляются определенные документы (отчеты), заполненные на основании данных первичного учета и подписанные лицами, отвечающими за достоверность содержащихся в них сведений о деятельности предприятий, организаций и учреждений в отчетном периоде. Данные отчетности служат основанием для составления плана и контроля за ходом его выполнения, для оперативного руководства, а также для всякого рода социально-экономических исследований. Органы государственной статистики должны совершенствовать систему показателей отчетности, устанавливать такие сроки ее разработки, чтобы всегда можно было иметь достаточно ясную картину процесса выполнения плана и на этой основе не только вовремя устранять недо-

статки, но и быстро использовать резервы роста производ-

ства. Отчетность в нашей стране весьма разнообразна и своими возможностями охватывает все основные процессы, характеризующие рост материально-технической базы коммунизма и внедрение в жизнь различных социально-культурных мероприятий. Но как бы ни был велик комплекс показателей отчетности, мы все же никогда не сможем получить все сведения, которые нужны для управления и планирования. К таким сведениям относятся, например, данные о составе населения (по полу, возрасту, образованию, общественным группам и т. п.), о размерах и структуре доходов и расходов семей рабочих и колхозников, о количестве и составе товаров на колхозных рынках и т. п. Поэтому советская статистика в довольно широких масштабах проводит специально организованное статистическое наблюдение в форме переписей и всякого рода обследований.

Полученные в результате проведения специально организованного статистического наблюдения сведения часто используются для уточнения и проверки данных отчетности, а также для более глубокого и всестороннего анализа общественных явлений, в частности при всякого рода социологических исследованиях.

По частоте регистрации фактов статистическое наблюдение подразделяется на непрерывное и прерывное.

Под непрерывным, или текущим, наблюдением понимается такое наблюдение, при котором установление и регистрация фактов производится по мере их возникновения. При непрерывном наблюдении регистрация фактов сопутствует течению событий и поэтому производится постоянно. Так производятся, например, регистрация актов гражданского состояния (рождения и смерти, браки и разводы), учет выпуска продукции, отгрузки и прибытия грузов и т. п.

Прерывным называется такое наблюдение, при котором регистрация фактов производится либо регулярно через определенные промежутки времени, либо по мере надобности. Если наблюдение проводится регулярно через равные промежутки времени, то оно называется периодическим (ежегодные переписи остатков сырья и материалов у потребителей, ежегодный заключительный учет итогов сева под урожай текущего сельскохозяйственного года и т. п.). Если же наблюдение проводится от случая к случаю, т. е. по мере надобности, то оно называется единовременным. Примерами единовременного статистического наблюдения в СССР могут служить переписи многолетних насаждений, переписи жилищного фонда и др.

При выборе вида наблюдения в каждом конкретном случае исходят из характера изучаемого явления и особенностей его развития, из большей или меньшей потребности в соответствующих сведениях, а также из соображений экономической целесообразности более или менее частого проведения наблюдения. Как пра-

вило, для точного учета большинства процессов общественной жизни необходимо текущее наблюдение (учет естественного движения населения, учет производства продукции и т. п.). Однако ограничиться одной только текущей статистикой нельзя, так как, во-первых, в ходе текущего наблюдения могут постепенно накапливаться ошибки, а, во-вторых, текущее наблюдение не может охватить все многочисленные процессы общественной жизни, так как это потребовало бы очень больших затрат времени и средств. Поэтому наряду с текущим наблюдением прибегают к проведению периодических и единовременных обследований. Так, например, отчетность промышленных предприятий об остатках, поступлении и расходе сырья и материалов (форма № 1-СН) дает возможность следить за изменениями в количестве и ассортименте материалов. Тем не менее проводятся переписи материальных ресурсов. Задача таких переписей заключается в том, чтобы учесть остатки по более широкой номенклатуре и более широкому кругу потребителей.

Данные переписей часто служат для разного рода исчислений. Так, изменения в численности и размещении населения по территории страны можно изучить на основании данных текущей статистики населения. Но для того чтобы определить численность населения на определенную дату, необходимо к количеству населения по последней переписи прибавить естественный прирост и сальдо миграции (т. е. перемещения населения) по данным текущего учета за период, прошедший с момента переписи.

По охвату единиц обследуемой совокупности различают сплошное и несплошное наблюдение.

Сплошным называется наблюдение, при котором учету подлежат все единицы изучаемой совокупности. Примерами сплошного наблюдения являются переписи населения, переписи плодовых насаждений и виноградников и т. д. Текущая статистика, базирующаяся на отчетности, также относится к сплошному виду наблюдения, так как она с целью контроля за выполнением государственного плана охватывает все подотчетные единицы.

По ряду причин (в целях экономии времени и средств, в связи с невозможностью сплошного наблюдения из-за уничтожения или порчи обследуемых единиц и т. п.) в широких масштабах проводятся также несплошные наблюдения. Несплошным называется наблюдение, при котором регистрации подлежат не все единицы изучаемой совокупности. Оно подразделяется на следующие виды: наблюдение основного массива, анкетное, монографическое и выборочное наблюдение.

Наблюдение основного массива представляет собой такой вид несплошного наблюдения, при котором из всей совокупности единиц объекта наблюдению подвергается такая их часть, у которой объем изучаемого признака составляет подавляющую, преобладающую долю в общем его объеме. Проведение наблюдения по этому методу практикуется в тех случаях, когда сплошной охват единицы совокупности сопряжен с особыми трудностями и в то же

время исключение из наблюдения определенного количества единиц не оказывает существенного влияния на величину обобщающих показателей. С его помощью производится, например, изучение цен и количества товаров, проданных на колхозных рынках. Так, регистрация цен колхозной торговли осуществляется крупнейшими органами государственной статистики в 264 наиболее крупных городах, доля которых в общем объеме колхозной торговли страны весьма значительна. При этом в каждом городе учет цен производится не на всех рынках, а только на основных, занимающих наибольшую долю в товарообороте колхозной торговли данного города. Неполнота охвата городов и колхозных рынков на результаты наблюдения существенно повлиять не может.

Анкетный вид наблюдения сводится к тому, что необходимые данные получаются путем рассылки специальных вопросников (анкет), заполнение и возвращение которых основано на добровольных началах. Поэтому фактическое количество полученных анкетных ответов обычно меньше числа разосланных.

В нашей стране этот метод используется органами связи для наблюдения своевременности доставки писем и газет, издательскими организациями для получения отзывов о книгах и т. п. Анкетный вид наблюдения был использован Институтом общественного мнения, который с помощью газеты «Комсомольская правда» провел много анкетных обследований с целью получения представления о суждениях, мнениях и жизни советской молодежи. В последнее время анкетное наблюдение применяется различными учреждениями и организациями при проведении конкретных социологических исследований.

Наиболее совершенным видом несплошного наблюдения является выборочное. Выборочным называется такой вид несплошного наблюдения, при котором с целью характеристики всей совокупности обследованию подвергается некоторая ее часть, взятая в выборку по определенным правилам. При этом уже не ставится задача охвата основного массива. Основным условием правильности проведения выборочного наблюдения является такой отбор, в результате которого отобранная часть единиц по всем подлежащим изучению признакам достаточно точно характеризовала бы всю совокупность в целом.

В нашей стране в практике статистической работы выборочное наблюдение с каждым годом приобретает все большее распространение.

Особым видом несплошного наблюдения являются монографические описания, которые заключаются в подробном описании отдельных типичных объектов: колхозов, районов, передовиков производства и т. п. С помощью монографического описания достигаются углубленное, детальное изучение вопросов, которые не могут быть освещены при массовом обследовании. Монографическим описаниям большое значение придавал В. И. Ленин. Он

Вопросы выборочного наблюдения рассматриваются в гл. VII.

требовал от советских статистиков и руководителей местных организаций сообщать точные сведения о числе коммуны, колхозов, артелей и других видов общественного хозяйства с подразделением на поставленные заведомо хорошо, сносно и неудовлетворительно. В. И. Ленин рекомендовал: «Одно типичное хозяйство каждой из этих последних трех групп должно быть не менее двух раз в год описываемо подробно с точным указанием всех данных об описываемом хозяйстве, его величии, места нахождения, итогов производства, помощи от него крестьянскому хозяйству и т. д.»¹.

Особенно большое значение монографическому виду наблюдения придается в последние годы. Всесоюзное совещание статистиков (1968 г.) в своем решении указало, что для дальнейшего совершенствования методов сбора статистической информации необходимо обеспечить «широкое внедрение... монографических статистических обследований по отдельным актуальным проблемам хозяйственного и культурного строительства».

Способы наблюдения. По способу осуществления наблюдения различают непосредственное наблюдение, документальный способ и опрос.

При непосредственном наблюдении представители органов государственной статистики или других организаций получают необходимые сведения путем личного подсчета, измерения, взвешивания и т. п. единиц объекта наблюдения. Примером такого способа наблюдения может служить перепись садов и огородников, при которой регистраторы имеют право записывать в переписные листы количество плодовых деревьев в каждом хозяйстве лишь после их пересчета, а не со слов владельца.

При документальном способе наблюдения необходимые данные получают на основании использования различной документации. Этот способ наблюдения часто называют отчетным, так как он применяется главным образом при заполнении предприятиями, организациями и учреждениями статистической отчетности на основе документов первичного учета. Типичным примером этого способа наблюдения может служить суточная, декадная, месячная отчетность предприятий о производстве чугуна, стали и т. д. Характерным примером документального способа наблюдения является также использование органами государственной статистики вторых экземпляров актов гражданского состояния для получения сведений о естественном движении населения.

Опрос может производиться в устной и письменной форме. При устном опросе специально подготовленные работники (регистраторы, счетчики) опрашивают соответствующих лиц и с их слов записывают в бланки нужные сведения. Этот способ наблюдения называется также экспедиционным. Характерным примером использования экспедиционного способа является Всесоюзная перепись населения 1970 г., когда счетчики заполняли переписные

листы на основе устного опроса граждан при обходе жилых и других помещений.

При письменном опросе опросные бланки раздаются или рассылаются опрашиваемым лицам, которые сами их заполняют. Представители же статистических органов лишь инструктируют опрашиваемых лиц и контролируют полноту и правильность сведений при получении заполненных бланков. Этот способ называют также саморегистрацией. Применяют его при бюджетных обследованиях рабочих, служащих и колхозников, а также при проведении некоторых переписей (например, перепись неустановленного оборудования на 1 января 1971 г. и др.).

Разновидностью письменного опроса является корреспондентский способ наблюдения. Он заключается в том, что необходимые сведения статистическим органам сообщают постоянные корреспонденты. Этот способ в нашей стране применялся до 1929 г. для изучения процессов, происходящих в сельском хозяйстве. В настоящее время для тех же целей корреспондентский способ широко применяется в Польской Народной Республике. Другая разновидность письменного опроса — анкетный опрос, при котором заполненные бланки в присутствии опрашиваемого не проводятся.

Выбор способа наблюдения зависит прежде всего от характера объектов статистического наблюдения, от предъявляемых требований к степени точности данных и, наконец, от финансовых возможностей и соображений о кадрах.

При выборе способа наблюдения необходимо иметь в виду преимущества и недостатки каждого из них. Наиболее совершенным способом получения необходимых данных является непосредственное наблюдение, которое при соблюдении всех правил статистической науки обеспечивает наибольшую достоверность статистических материалов. Но проведение наблюдения этим способом часто требует больших затрат для оплаты счетчиков, регистраторов и т. п. В социалистических странах при организации соответствующего контроля за состоянием первичного учета и достоверностью данных отчетности вторым по точности получаемых данных является документальный способ наблюдения. При проведении статистического наблюдения корреспондентским способом или путем саморегистрации достоверность данных меньше.

2. ПЛАН СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ

Содержание плана наблюдения. Всякое статистическое наблюдение проводится по заранее и детально разработанному плану. План статистического наблюдения содержит решение программно-методических и организационных вопросов. К программно-методическим вопросам относятся: установление цели и задач наблюдения, определение объекта и единицы наблюдения, разработка программы наблюдения, а также выбор вида и способа

¹ Ленин В. И. Полн. собр. соч., т. 43, с. 282—283.

наблюдения. К *организационным* — установление места, времени и сроков наблюдения, определение круга лиц и организаций, отвечающих за проведение наблюдений, подбор, обучение и инструктаж кадров, размножение и рассылка формуляров и другого инструментария наблюдений, установление сроков сдачи (представления) материалов, а также другие практические вопросы, связанные с подготовкой, организацией и проведением статистического наблюдения.

Конкретное решение всех программно-методологических и организационных вопросов определяется характером объекта исследования и поставленными перед ним целями и задачами.

Цель наблюдения

Установление цели и задач является исходным пунктом при организации любого статистического исследования. Нельзя приступить к статистическому наблюдению, не зная его цели. Цель наблюдения должна быть сформулирована ясно, совершенно четко и развернуто, с указанием конкретных задач, стоящих перед данным наблюдением. Например, цель Всесоюзной переписи населения 1970 г. заключалась в получении данных о численности населения, о его размещении по республикам, экономическим и административным районам, о составе населения по полу, возрасту, семейному положению, национальной принадлежности, уровню образования, а также распределения по источникам средств существования, занятиям физического и умственного труда, отраслям народного хозяйства и общественным группам. В отличие от предшествующей переписи при проведении переписи 1970 г. также ставилась цель — получить данные о продолжительности работы лиц, занятых в народном хозяйстве неполный год, составе лиц, занятых в домашнем и личном подсобном сельском хозяйстве, о миграции населения и др.

Руководящим методологическим центром, планирующим статистические работы в нашей стране, является ЦСУ СССР, в обязанности которого входит составление общесоюзного плана учетно-статистических работ. Привлечение к разработке этого плана широкого круга научных и практических работников обеспечивает продуманную организацию отчетности, а также проведение специально-организованных наблюдений.

В соответствии с целями и задачами определяются объект и единица статистического наблюдения. *Объектом наблюдения* называется совокупность социально-экономических явлений и процессов, которые подлежат исследованию.

Определить объект наблюдения — это значит установить, какая именно совокупность должна быть подвергнута наблюдению, определить ее границы и направления ее изучения. Для этого нужны глубокие знания об объекте.

Определение и отграничение объекта должны базироваться на всестороннем выяснении всех признаков, т. е. свойств и характерных особенностей, которые отличают исследуемое явление или

объект от других. Например, при переписи населения необходимо установить, какое именно население подлежит регистрации: наличное, т. е. фактически находящееся в данной местности в момент переписи, или постоянное, т. е. живущее в данной местности постоянно.

В ряде случаев для отграничения объекта наблюдения пользуются тем или иным ценом. *Под ценом* в данном случае понимается один или несколько заранее установленных признаков, определяющих границы объекта. Так, например, в Украинской ССР при переписи населения 1970 г. к городам относились благоустроенные населенные пункты с населением не менее 10 тыс. жителей, а к поселкам городского типа — пункты, насчитывающие более 2 тыс. жителей, среди которых не менее 60% составляли рабочие, служащие и члены их семей. Кроме того, к рабочим поселкам относились населенные пункты при промышленных предприятиях, стройках и станциях железной дороги с населением более 500 человек, преимущественное большинство которых — рабочие и служащие.

Определяя объект исследования, необходимо точно указать единицу изучаемой совокупности и единицу наблюдения.

Под единицей совокупности понимается составной первичный элемент объекта наблюдения, который служит основой счета и обладает признаками, подлежащими регистрации при проведении исследования. Следовательно, единица совокупности — это то, что при статистическом наблюдении будет подвергнуто обследованию. Из численности единиц совокупности определяют объем и распространенность изучаемого явления.

В отличие от единицы совокупности *под единицей наблюдения* понимают первичную ячейку, от которой должны быть получены необходимые сведения. Правильное определение единицы наблюдения важно при решении вопросов организации получения сведений.

Единицы совокупности иногда совпадают с единицами наблюдения. Например, при проведении специального обследования с целью изучения численности и состава лиц, занятых в домашнем или подсобном сельском хозяйстве, которое осуществлялось одновременно с переписью населения 1970 г., каждое лицо в трудовом возрасте, занятое на момент переписи в домашнем и личном подсобном сельском хозяйстве, являлось не только составным первичным элементом объекта обследования (единица совокупности), но и источником сведений (единица наблюдения). Но не всегда единица совокупности отличается от единицы наблюдения. Например, при проведении переписи плодово-ягодных насаждений единицей совокупности будет отдельное плодовое дерево (например, яблоня), ибо признаки, регистрируемые в этом обследовании (возраст насаждений, их породный состав и т. п.), относятся не к колхозу, совхозу и т. д., а к каждому плодovому дереву. Единицей наблюдения в этой переписи будут колхозы, совхозы,

организации, хозяйства колхозников, служащих и других групп населения, имеющих плодово-ягодные насаждения.

Четкое определение единицы совокупности и единицы наблюдения имеет большое значение. Оно дает возможность обеспечить не только полноту и точность учета всех элементов объекта обследования, но и правильность обработки и анализа результатов статистического наблюдения.

Неясность в этом вопросе приводит к путанице во всей дальнейшей работе. Критикуя дореволюционную правительственную фабрично-заводскую статистику, В. И. Ленин подчеркивал, что ее данными нельзя пользоваться без дополнительной их обработки именно вследствие того, что понятия «фабрика» и «завод» были в ней крайне неопределенны и расплывчаты, т. е. вследствие неточного отграничения объекта наблюдения и единицы совокупности¹.

При проведении статистического наблюдения в форме отчетности единицу наблюдения называют отчетной единицей, понимая под этим предприятия, учреждения и организации, которые обязаны представлять статистическую отчетность и несущие ответственность за достоверность ее данных. В промышленности отчетными единицами обычно являются заводы, фабрики, шахты и т. п., а в сельском хозяйстве — совхозы, колхозы и т. д.

Исходя из содержания объекта, цели и задач статистического исследования разрабатывается программа наблюдения.

Программа наблюдения

Программа наблюдения представляет собой перечень признаков и показателей, подлежащих регистрации. Иными словами, программа — это перечень вопросов на которые должны быть получены ответы по каждой единице наблюдения.

Программа статистического наблюдения должна быть построена на научной основе. В нее нужно включать такие вопросы и признаки, которые имеют наибольшее практическое и теоретическое значение и являются наиболее существенными для данного объекта.

Разработка программы наблюдения — одна из наиболее важных и ответственных задач. К ее содержанию предъявляются большие требования. Основными из них являются следующие.

В программу должны включаться только те признаки, значения которых неодинаковы у всех единиц изучаемой совокупности.

В программу следует включать только те вопросы, которые, безусловно, необходимы и ответы на которые будут использованы или при разработке материалов наблюдения, или в контрольных целях.

Программу наблюдения целесообразно строить так, чтобы ответами на одни вопросы можно было контролировать ответы на другие; например, в годовом отчете колхоза наряду с вопросом о размерах посевных площадей под каждой культурой имеется также вопрос и об общей площади посева.

Вопросы программы должны быть сформулированы четко, кратко и ясно. Требуется сформулировать вопросы так, чтобы они были поняты одинаково, а ответы на них можно было бы сравнить между собой.

Программа наблюдения должна разрабатываться коллективно, с привлечением практических и научных работников, которые глубоко знают существо явлений, подлежащих статистическому исследованию. Она должна быть всесторонне обсуждена в прессе, на совещаниях. В. И. Ленин в 1898 г. писал, что один из основных недостатков тогдашней фабрично-заводской статистики была полная неразработанность программы собирания сведений. «Если такая программа, — продолжал В. И. Ленин, — будет выработана в канцеляриях, не подвергаясь критике специалистов и (это особенно важно) всестороннему обсуждению прессы, то сведения никогда не могут быть сколько-нибудь полными и единообразными»¹. В. И. Ленин отмечал также, что «правильная программа и обеспечение проверки данных, это — два важнейшие условия успешной статистики»².

Статистические формуляры

Статистическим формуляром называется особый документ, в котором регистрируются ответы на вопросы программы. Различают две системы формуляров — карточную и списочную.

При карточной системе каждая карточка-формуляр предназначается для регистрации одной единицы наблюдения и ее признаков.

При списочной системе в одном формуляре регистрируются сведения о нескольких единицах наблюдения. Примером списочной системы формуляров могут служить переписные листы переписи населения СССР в 1970 г., каждый из которых был предназначен для регистрации шести лиц и характеризующих их признаков.

В настоящее время, когда результаты большинства статистических наблюдений разрабатываются на вычислительных машинах, списочная система формуляров имеет ряд преимуществ перед карточной. При машинной обработке все признаки, характеризующие единицу наблюдения, должны быть перенесены на особые разработанные бланки или перфорационные карточки, а эту работу пронумеровать со списка гораздо удобнее и быстрее, чем с карточек. Кроме этого, при проведении наблюдения на формулярах и в виде списка достигается экономия бумаги, более быстрая и тщательная проверка и т. д. Но в ряде случаев карточная система по сравнению со списочной имеет свои преимущества: в формуляр можно включать гораздо большее количество признаков; при разработке итогов наблюдения можно непосредственно подсчитывать частоты, располагая их по группировочным признакам. Вслед-

¹ См.: Ленин В. И. Полн. собр. соч., т. 4, с. 30.

² Ленин В. И. Полн. собр. соч., т. 4, с. 30.
там же, с. 4.

стве этих достоинств, а также в связи с тем, что в нашей статистике основными источниками получения статистических данных является отчетность, в настоящее время в практике советской статистики преимущественно используется картонная система формуляров.

Обязательным атрибутом каждого формуляра является адресная часть, в которой указывается, где и когда собраны сведения о данной единице наблюдения. Адресная часть используется главным образом в контрольных целях и для получения сведений в территориальном разрезе, а иногда и для получения каких-либо дополнительных сведений.

Для обеспечения единообразного толкования вопросов программы и облегчения их понимания в статистических формулярах после вопроса часто в виде подсказа перечисляется либо все, либо некоторые из возможных ответов или же даются указания о способе расчета того или иного показателя. Так, в переписном листе переписи населения 1970 г. подобного рода подсказы были сделаны к большинству вопросов. В частности, после первого вопроса: отношение к главе семьи — в скобках указано: жена, муж, сын, дочь, мать, отец, сестра, племянник, зять, свекор, теща и т. п. (см. приложение).

Подсказы содержатся также в большинстве форм отчетности. Так, в годовом отчете колхоза за 1972 г. в табл. 5 «Площадь и сбор урожая сельскохозяйственных культур в 1972 году» указывается, что для получения данных гр. 3 «Средний сбор с 1 га — ц» необходимо гр. 2 «Валовой сбор в первоначальном оприходованном весе (ц)» разделить на гр. 1 «Фактически посеяно (по озимым указать площадь, сохранившуюся к окончанию сева яровых) — га» (см. также табл. 3.1 на с. 70—71).

Как бы хорошо ни были сформулированы вопросы программы, для обеспечения более тщательного и правильного выполнения всех требований обычно к программе наблюдения составляются специальные письменные указания и разъяснения — инструкции.

В них излагают вопросы, связанные как с программно-методологической, так и с планово-организационной сторонами статистического наблюдения, а именно: о цели и задачах статистического исследования; об объекте и единице наблюдения; об отчетной единице; о времени и сроках проведения наблюдения; об оформлении материалов наблюдения и сроках их представления в вышестоящие организации. Независимо от разъяснений в самом опросном бланке в инструкциях обычно дают указания, как следует понимать тот или другой вопрос программы, иногда указывают примерные ответы, дают типичные случаи заполнения формуляров.

Например, в указаниях о записи ответа на вопросы переписного листа переписи населения 1970 г. к ятяму вопросу «Возраст» были даны такие пояснения: «Для лиц старше одного года записываются цифрами только число исполнившихся лет (например, лицам, которым к моменту переписи исполнилось

27 лет и 11 месяцев или 27 лет и 1 месяц, записывается 27 лет). Для детей старше одного года записывается только число исполнившихся месяцев. Для детей младше одного месяца записывается «мнеся месяца».

Понят селся записывать с округлением (например, 39-летним нельзя указать, что им 40 лет, 36-летним — что им 35 лет и т. д.; лицам, которым исполнилось 22 года и 10 месяцев, нельзя округлять возраст до 23 лет). Сметчик указывал, кроме того, выяснить и записать месяц и год рождения и установить, что возраст округляемого указан точно без округления. При этом месяц и год рождения записываются цифрами (например, IV 1958).

Изложение инструкции должно быть ясным, простым, четким и вместе с тем кратким. При разработке ее следует достаточно ясно и подробно прокомментировать все детали программы и дать лицам, проводящим наблюдение, необходимые указания так, чтобы все основные вопросы, которые могут возникнуть в процессе наблюдения, получили соответствующее разъяснение.

Для обеспечения успеха статистического наблюдения важно правильно решить не только программно-методологические, но и организационные вопросы. **Перечень мероприятий, необходимых для успешного проведения работы по сбору сведений, называется организационным планом статистического наблюдения.**

Среди вопросов организационного плана важное значение имеет точное установление времени, к которому относятся регистрируемые сведения. Необходимо также установить и время, в течение которого будет проводиться наблюдение. В связи с этим различают объективное и субъективное время наблюдения.

— **Время, к которому относятся регистрируемые сведения, называется объективным временем наблюдения.** Оно представляет собой либо определенный момент времени, либо тот или иной его период. Например, сведения о выпуске продукции, о размере товарооборота и т. п. всегда относятся к тому или иному периоду времени: суткам, месяцу, кварталу, году. Итоги любого процесса нельзя подсчитать иначе, как за какой-либо период времени. Сведения же о наличии каких-либо единиц, например о численности населения, о размере имеющегося жилищного фонда и т. п., могут быть отнесены только к определенному моменту времени, по состоянию на который произведена регистрация. **Момент времени (день, час), к которому приурочены сведения, называется критическим моментом наблюдения.** Так, критическим моментом при переписи населения 1970 г. было 12 часов ночи с 14 на 15 января. Поэтому лица, умершие до 12 часов ночи, регистрации не подлежали, а умершие после 12 часов ночи вносились в переписные листы, так как на критический момент переписи они еще были живы. Родившиеся до 12 часов ночи записывались в переписные листы, а родившиеся после 12 часов регистрации не подлежали. В результате установления критического момента итоги переписи являются как бы фотоснимком состояния населения на этот критический момент.

При статистическом наблюдении в форме отчетности объективным временем может быть как момент (дата), так и период.

Например, в ежемесячном отчете колхозов и совхозов о производстве продукции животноводства (форма № 24-сх) численность скота и птицы показывается по состоянию на 1-е число каждого месяца, а сведения о полученном приплоде, падеже, надое молока и т. п. — за период с начала года.

С объективным временем не следует смешивать субъективное время наблюдения, т. е. тот период, в течение которого производится регистрация. При проведении переписей обычно устанавливаются даты начала и окончания наблюдения, т. е. его календарные сроки. Так, перепись населения 1970 г. продолжалась 8 дней (с 15 по 22 января).

Место и органы наблюдения

При планировании наблюдения нужно установить, в пределах какой территории оно будет проводиться, а также место, где непосредственно будет осуществляться регистрация. Установление места проведения наблюдения имеет важное значение. Место проведения наблюдения часто совпадает с местом нахождения единицы наблюдения. Так, отчетность по тому или иному вопросу о деятельности колхоза, промышленного предприятия составляется по месту их нахождения. Однако в ряде случаев необходимо точно указать, где, в каком именно месте должно быть проведено наблюдение. Например, при проведении Всесоюзной переписи населения 1970 г. в инструкции указывалось, что население должно учитываться по месту жительства, а не по месту работы.

При документальном способе вопрос о месте наблюдения должен быть решен для того, чтобы точно определить, где должны быть документально оформлены наблюдаемые или учитываемые факты, где должен проходить, например, табельный учет — в пределах отдельных цехов или он должен быть централизован и т. п.

В плане наблюдения должно быть также предусмотрено, кто проводит наблюдение. В СССР в масштабе всей страны статистическое наблюдение организуют ЦСУ СССР и его местные органы, министерства, ведомства в лице их учетно-статистических отделов в центре и на местах, общественные организации и т. д. В плане четко определяются права и обязанности отдельных учреждений и организаций, принимающих участие в проведении данного наблюдения, а также указывается, что должны делать министерства, ведомства, права и обязанности ЦСУ СССР и его местных органов и т. д.

Подготовительные работы

Успешное проведение статистического наблюдения немисливо без осуществления ряда подготовительных мероприятий. Особенно большое внимание подготовительным мероприятиям уделяется при разработке организационного плана проведения переписей, которые представляют собой особо сложный вид статистического наблюдения. Одной из первых подготовительных работ является составление списков всех единиц, которые должны представлять сведения. Так, при проведении учета посе-

вых площадей инспектуры государственной статистики составляются списки колхозов, совхозов и организаций, находящихся на территории района (города), а также списки населенных мест с указанием количества хозяйств колхозников и других групп населения.

Следующий этап подготовительной работы — разработка территории на участки и расчет необходимого количества кадров. Расчет количества переписных кадров находится в тесной связи с нормой нагрузки и количеством единиц наблюдения и единиц совокупности, которые устанавливаются на основании списка населенных мест, домовладений, списка промышленных предприятий, колхозов, совхозов и т. д. Чем больше единиц совокупности, тем, разумеется, должно быть больше регистраторов и другого персонала. Численность кадров, привлекаемых для проведения наблюдения, зависит также от срока производства наблюдения, от программы наблюдения, способа сбора материалов, от объема работ по распределению единиц наблюдения и единиц совокупности по территории и т. п. При разработке территории на участки исходят из количества регистраторов и нормы их нагрузки; принимаются также в расчет разбросанность единиц и транспортные условия.

Большое внимание при проведении подготовительных работ должно быть уделено подготовке кадров, которая сводится к инструкторскому и обучению переписных кадров, к практическому ознакомлению ими переписных формуляров и т. д.

К числу важнейших подготовительных работ к переписям относятся и пропаганда. В нашей стране статистические обследования воспринимаются населением как всенародное дело и поэтому находят всемерную поддержку и содействие. Особенно ярко это проявилось при проведении переписей населения в 1939, 1959 и 1970 гг.

При подготовке к проведению наблюдения необходимо своевременно отпечатать бланки, инструкции и другую документацию и разослать их на места.

В подготовительные работы включается также проведение пробных переписей. Пробные обследования обычно проводятся по тому же плану и программе, что и основное. Задачи этих переписей заключаются в проверке на практике проекта программы и организационного плана наблюдения. Так, для обеспечения успешного проведения Всесоюзной переписи населения 1970 г. в одном из районов была проведена пробная перепись населения в 9 районах страны, отличающихся природно-географическими и экономическими условиями, национальным составом и характером населения.

Пробное обследование уместно и тогда, когда формой наблюдения является отчетность. В данном случае оно производится для того, чтобы проверить, насколько данный комплекс показателей удовлетворяет разрешенной задаче, поставленной перед статистическим исследованием.

4. ПРАКТИКА ПРОВЕДЕНИЯ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ В СССР

Первичный учет и отчетность

Правильной организацией системы отчетности в нашей стране всегда придавалось большое значение. В. И. Ленин писал: «Поставка отчетности, например, есть вещь основная во всех ведомствах и учреждениях самых разнородных»¹.

Согласно указаниям В. И. Ленина, на всех этапах экономической политики КПСС организации системы отчетности, ее совершенствованию уделялось большое внимание.

После XXIV съезда КПСС в связи с необходимостью улучшения системы управления народным хозяйством, планирования и стимулирования производства подвергаются пересмотру содержания форм статистической отчетности и сроки ее представления. Отчетность должна быть простой, но содержать такой объем показателей, который позволял бы всесторонне изучать и анализировать результаты деятельности каждого предприятия и производить своевременную корректировку планов в процессе их выполнения.

Отчетность предприятий, учреждений и организаций в СССР составляется на основании данных первичного учета. *Первичным учетом называется ведение систематических записей в формах первичных учетных документов о различных фактах, относящихся к деятельности предприятий, учреждений и организаций.* Формы первичных учетных документов в своем большинстве используются для составления статистической отчетности. Например, при выполнении производственного плана (форма № 1-п) используются такие первичные документы, как ведомость выпуска продукции и накладные на прием и передачу готовых изделий на склад и на отпуск со склада.

В документах первичного учета отражается вся деятельность предприятий, учреждений и организаций и результаты их работы. Поэтому состояние первичного учета непосредственно определяет качество отчетности.

В Директивах XXIV съезда КПСС указано на необходимость «всемерно улучшить систему учета и отчетности, совершенствовать статистику». Эта задача направлена на то, чтобы все виды учета отвечали возросшим требованиям руководства развитием народного хозяйства. В период развернутого строительства коммунизма нужна такая отчетность, которая позволяла бы во всей полноте, а главное своевременно реагировать на ход выполнения плана, распространять передовой опыт и устранять допущенные ошибки.

В последние годы проводится большая работа по унификации первичной учетной документации и ее внедрению. Унификация предполагает, что однородные предприятия, отрасли должны применять единые первичные учетные документы для оформления

однородных хозяйственных операций. Это позволяет концентрировать в одном бланке необходимые показатели для учета, использовать единые учетные шифры и коды, упростить формы и приспособить их к заполнению и обработке на вычислительных машинах.

В свете решений XXIV съезда КПСС работа по унификации первичной учетной документации должна быть усилена, чтобы обеспечить более широкое применение типовых форм учета и создать условия для полной механизации учета и отчетности.

Выполняя указания XXIV съезда КПСС, все предприятия и организации страны независимо от их подчиненности со второго полугодия 1973 г. перешли на утвержденные ЦСУ СССР новые типовые межведомственные формы первичной документации по учету извещений, основных средств, сырья и материалов, капитального строительства, перевозок автомобильным транспортом, рационализации и изобретательства, заготовок сельскохозяйственных продуктов и сырья. Ведется работа по подготовке типовых форм первичной учетной документации для учета и других сторон деятельности предприятий и организаций.

Составляемая предприятиями, учреждениями и организациями отчетность представляется вышестоящим организациям, органам ЦСУ СССР и другим организациям в порядке, который устанавливает ЦСУ СССР по каждой форме отчетности.

В СССР различается общегосударственная и внутрисоветовская отчетность. *Общегосударственной является отчетность, обязательная для всех предприятий, учреждений и организаций и представляемая в сводном виде непосредственно через ЦСУ СССР правительству.*

Внутрисоветовской является отчетность, получаемая министерствами и ведомствами для своих оперативных нужд.

Формы государственной отчетности могут быть типовыми и специализированными. Типовые формы содержат одни и те же показатели для всех предприятий, учреждений и организаций данной отрасли народного хозяйства или всех отраслей народного хозяйства. В специализированной отчетности вопросы видоизменяются в зависимости от особенностей отдельных отраслей народного хозяйства.

Все формы статистической отчетности утверждают органы государственной статистики. ЦСУ СССР утверждает общесоюзный минимум показателей и формы статистической отчетности для предприятий, строек и организаций, подчиненных общесоюзным и союзно-республиканским министерствам и ведомствам. ЦСУ союзных республик имеют право исходя из экономических особенностей отдельных районов и с учетом утвержденного ЦСУ СССР общесоюзного минимума показателей статистической отчетности устанавливать при необходимости дополнительную отчетность для предприятий и организаций союзного и союзно-республиканского подчинения. ЦСУ союзных республик утверждают также объем и формы статистической отчетности для предприятий, строек, учреждений и организаций местного подчинения.

¹ Ленин В. И. Полн. собр. соч., т. 44, с. 127.

Отчетность по формам, не утвержденным ЦСУ СССР или ЦСУ союзных республик, является незаконной. Правительство запрещает всем государственным органам и общественным организациям требовать, а руководители предприятий, учреждений и колхозов заполнять и представлять отчетность по формам и показателям, не утвержденным ЦСУ. Органам государственной статистики предоставлено право отменять незаконную отчетность, а лиц, виновных во введении такой отчетности, привлекать к ответственности.

В каждой форме утвержденной отчетности предусматриваются следующие обязательные реквизиты: а) дата и номер утверждения формы ЦСУ СССР или ЦСУ союзной республики; б) название формы, определяющее содержание отчетности; в) период, за который представляется отчет, или дата, к которой относятся отчетные данные; г) срок представления отчетности, дата высылки или через сколько дней после отчетного периода она высылается; д) адреса, в которые должна представляться отчетность; е) номер (или литеры), присвоенный данной форме; ж) название предприятия, учреждения или организации, которое представляет отчетность, а также название государственного комитета, министерства (ведомства), треста, комбината, управления, которым подчинено предприятие; з) местонахождение (республика, край, область) и адрес предприятия или учреждения, представляющего отчет; и) должности лиц, обязанных подписать отчет и ответственных за его составление.

Система отчетности и ее объем определяются перечнем (табелем) отчетности, утвержденным ЦСУ СССР. Табеля представляют собой список действующих форм оперативно-статистической отчетности. Табеля содержат следующие сведения: наименование и номера форм отчетности, способ представления (почтой, телеграфом), кто представляет отчетность (предприятие, комбинат, трест и т. д.), кому она представляется, периодичность отчетности, срок высылки.

На основании перечней отчетности ЦСУ СССР ежегодно составляет план статистических работ, выполняемых органами государственной статистики. В плане указываются для каждого звена статистических органов сроки окончания той или иной работы, выполняемой на основе поступающей отчетности. Таким образом, вся статистическая работа в нашей стране ведется по единому плану.

Предприятия, организации и учреждения представляют отчетность в органы государственной статистики и в свои вышестоящие организации в различные сроки: ежедневно, за неделю, за пятнадцать дней, за месяц, квартал, за год. Отчетность, представляемая в течение года, называется текущей, а за год — годовой.

Текущая отчетность в СССР представляет собой очень обширный статистический материал. Чем длительнее период, за который учитывается предприятие, организация или учреждение, тем обширнее программа отчетности. Ежедневная отчетность касается ограниченного круга важнейших показателей (производство ме-

талов, добыча угля и нефти и некоторые другие). Месячная отчетность освещает значительно больший круг вопросов, а годовой отчет еще полнее охватывает деятельность предприятий, учреждений или организаций, характеризует различные стороны их работы.

Так, например, в месячной отчетности промышленных предприятий по форме № 1-п содержится такие показатели (по плану и фактически за отчетный месяц и квартал, с начала отчетного года и за соответствующий период с начала прошлого года): объем реализованной и валовой продукции в оптовых ценах предприятий; производство продукции в натуральном выражении. В годовой отчет промышленного предприятия входят все показатели, характеризующие выполнение плана отчетного года. Кроме того, в годовом отчете содержится сведения о состоянии оборудования и сылового хозяйства предприятий, об объеме в составе основных фондов, об издержках производства и себестоимости продукции и др.

По способу представления отчетных данных различается телеграфная и почтовая отчетность. Телеграфная отчетность используется для получения сведений за самые короткие сроки, по особым важным вопросам: она содержит минимум показателей, обычно запрашиваемых на оперативный контроль за выполнением плана.

Система государственной отчетности в нашей стране постоянно развивается и улучшается. В целях совершенствования учета и статистики, а также в связи с необходимостью ускорения механизации учетно-статистических работ в 1957 г. проведена централизация собирания и разработки отчетности по промышленности, строительству, сельскому хозяйству в статистических органах (на машиноисчислительных станциях и в вычислительных центрах). Это привело к повышению качества и сокращению сроков получения сводных итоговых данных.

В результате централизации отчетности установлен следующий порядок ее представления:

1. Промышленные предприятия и стройки представляют отчетность непосредственно в областные, краевые и республиканские (АСР) статистические управления. Колхозы, совхозы и заготовительные организации представляют отчетность районным информационно-вычислительным станциям государственной статистики, которые ее контролируют, разрабатывают и представляют в сводном виде статистическим управлениям области (края), автономной республики, а также районным руководящим и плановым органам.

2. Статистические управления областей (краев) и автономных республик представляют сводную отчетность местным руководящим и плановым органам, а также ЦСУ союзных республик.

3. Центральные статистические управления союзных республик представляют сводные данные республиканским руководящим и плановым органам, а также в ЦСУ СССР.

ЦСУ СССР представляет сводную статистическую информацию Совету Министров СССР, Госплану СССР и другим центральным органам.

Централизация сбора и обработки отчетности обеспечивает укреплению связи органов государственной статистики с местными руководящими органами, улучшает практическое использование статистической информации для контроля за ходом выполнения плана.

Централизация учета и статистики создает все необходимые условия для скорейшего завершения механизации и автоматизации учетно-статистических работ, а следовательно, и коренного изменения порядка сбора и разработки статистической информации.

Решающую роль в изменении порядка сбора и обработки отчетности будут играть *автоматизированные системы управления (АСУ)*, которые начали создаваться в 1969 г. на предприятиях и в министерствах. За восьмью пятилетку внедрено 414 автоматизированных систем управления предприятиями и объединениями, создано 686 вычислительных центров (ВЦ). В девятой пятилетке (1971—1975 гг.) предусматривается создать и ввести в строй полнотные или в объеме первых очередей свыше 2700 автоматизированных или автоматических систем управления различного назначения.

В зависимости от комплекса применяемых на предприятии средств вычислительной и периферийной техники операции, связанные со сбором и обработкой информации, а также логические операции, связанные с анализом данных и выработкой решений, могут выполняться с различной полнотой и степенью механизации или автоматизации. Если операции по сбору и обработке информации, нужной для управления предприятием, выполняются в определенной непрерывности и взаимосвязи, то в этом случае действует автоматизированная система сбора и обработки информации (АССОИ.) Если же кроме этих выполняется еще и ряд логических операций, связанных с принятием решений, то в этом случае функционирует высшая ступень организации производства, которую называют автоматизированной системой управления предприятием (АСУП). В настоящее время уже действуют на Львовском телевизионном заводе АСУП «Львов», на Донецком машиностроительном заводе им. Ленинского комсомола Украины «Донецк» и др. АСУП включает в себя несколько подсистем. Одной из них является *подсистема отчетности и экономического анализа деятельности предприятия*. Эта подсистема обеспечивает составление установленной отчетности и решает задачи комплексного анализа всех показателей деятельности предприятия. Кроме этого в АСУП имеются еще три подсистемы: сбора и передачи информации, обработки информации, а также размножения, хранения и поиска информации. Эти подсистемы должны обеспечивать сбор информации на машинных носителях, ее обработку и хранение.

Первичный учет производственно-хозяйственной деятельности предприятия осуществляется в цехах, складах, отделах и службах с помощью автоматических устройств. Фиксирование первичной информации на машинный носитель (перфокарту, перфокарту-ленту, магнитную ленту и т. п.) осуществляется с помощью спе-

циального «машинного» языка — шифров, кодов. Зафиксированная на машинных носителях первичная информация после ее обработки на ВЦ предприятия используется для целей управления и является исходной для вышестоящей организации (объединение, главк, министерство), а также для вычислительных центров органов государственного управления (ЦСУ, Министерство финансов и др.).

Предусмотренное Директивами XXIV съезда КПСС по пятилетнему плану развития народного хозяйства СССР на 1971—1975 гг. дальнейшее развертывание работ по созданию и внедрению *автоматизированных систем планирования и управления отраслями (ОАСУ)*, объединениями, предприятиями должно привести к созданию *общегосударственной автоматизированной системы сбора и обработки информации для учета, планирования и управления народным хозяйством (ОГАС)*. В составе ОГАС будет действовать *автоматизированная система государственной статистики (АСГС)*.

АСГС будет создана на основе уже действующей и развивающейся сети вычислительных центров (ВЦ) и машиносчетных станций (МСС) ЦСУ СССР, имеющих большое количество ЭВМ. Ввиду того, что АСГС будет осуществлять сбор и обработку всей статистической информации, необходимой руководящим и плановым органам для решения задач по руководству страной, она будет являться одним из важнейших информационных каналов в системе управления народным хозяйством.

При автоматизированной системе управления на предприятиях и в организациях отчетность по заранее разработанным программам будет собираться на ЭВМ и фиксироваться на технических носителях. Зафиксированная на магнитных лентах, магнитных дисках и т. д. отчетность по каналам связи будет передаваться в ВЦ АСГС. Полученная ВЦ АСГС отчетность, после ее разбора в любом нужном направлении и разрезе, поступит в подсистему размножения, хранения и поиска статистической информации (АВД — автоматизированный банк данных).

Переписи и их организации

Наряду с отчетностью в нашей стране большое внимание уделяется проведению специально организованных переписей, учетов, обследований.

В настоящее время в СССР ежегодно проводят учет посевных площадей, скота, тракторов, переписи остатков материалов; через бумажные длительные промежутки времени — переписи населения, садов и огородников, учреждений здравоохранения, школьные переписи и т. д.

Переписи проводятся или путем использования данных учета и отчетности предприятий, учреждений и организаций или же путем специально организованной регистрации. К первому типу принадлежат переписи фактических остатков материалов: черных и цветных металлов, металлических и кабельных изделий, труб и т. д. Эти переписи производятся на основе данных бухгалтерского учета или данных инвентаризации наличных остатков материалов,

сверенных с данными бухгалтерского учета. По-другому проводятся переписи населения, садов и ягодников, торговых предприятий, оборудования и т. д. Так, перепись садов и ягодников у населения не может основываться на данных какого-либо систематически ведущего учета и поэтому проводится специально подготовленными счетчиками (регистраторами) путем обхода дворов.

Исходя из многолетней практики проведения переписей статистической наукой выработаны правила и принципы организации их проведения. Важнейшими из них являются следующие. Критический момент и сроки проведения переписи должны быть одинаковыми для всех единиц и для всей территории. Период проведения переписи должен наилучшим образом соответствовать ее задачам. Переписи следует проводить в возможно сжатые сроки. Целесообразно проводить переписи регулярно (желательно через одинаковые промежутки времени).

Классическим примером специально организованного статистического наблюдения являются переписи населения. Рассмотрим организацию и проведения переписей на примере *Всесоюзной переписи населения 1970 г.*

Успех переписи был обеспечен проведенной большой подготовительной работой. Прежде всего большое внимание было уделено разработке программы переписи. Разработанные заранее квалифицированными специалистами статистики проект программы переписи и основные вопросы организационного плана ее проведения были детально обсуждены на Всесоюзном совещании статистиков (апрель 1968 г.), в работе которого приняли участие не только статистики, но и другие специалисты.

В установленные организационным планом сроки были уточнены перечень и границы городских поселений, названия улиц, а также составлены списки домовладений в городских поселениях и списки сельских населенных пунктов. На основании этого территория всех районов и городов была разбита на переписные отделы, инструкторские и счетные участки. В переписные отделы включались в среднем: в городе — 8, а в сельских местностях — 6 инструкторских участков.

Особое большое внимание было уделено подбору и подготовке переписных кадров. К производству переписи было привлечено около 700 тыс. переписных работников и работников органов государственной статистики, в том числе до 550 тыс. счетчиков и 100 тыс. инструкторов-контролеров.

Огромное внимание было уделено массово-политической работе среди населения. Всеми средствами агитационной работы (лекции, печать, кино, телевидение и т. д.) трудящимся были разъяснены цели, задачи, политическое и народнохозяйственное значение переписи и порядок ее проведения. Переписи были восприняты населением и проведена как несложное дело.

Переписи населения в 1970 г. была проведена путем опроса населения по месту его жительства с 15 по 22 января. Результаты опроса записывались в переписные листы синюшечной формы, содержащие адресные сведения и программу переписи.

В отличие от всех предыдущих переписей, перепись населения 1970 г. проводилась по двум программам: 1) *программе сплошной переписи*, содержавшей 11 вопросов, и 2) *программе выборочной переписи*, имеющей дополнительно 7 вопросов. На вопросах выборочной переписи давали ответы все граждане. По проживающему населению. Общее спрашиваемых производилось механически. За типичный дом, изба, комната и общежитие и т. д. Переписывалось постоянно проживающее население (включая временно отсутствующих) каждого четвертого помещения, отмеченного инструктором-контролером до начала переписи в записной книжке счетчика.

Перепись населения 1970 г. была проведена по состоянию на 15 января. Выбор года, месяца и дня проведения переписи не является случайным. Проведение переписи именно в 1970 г. было обусловлено, во-первых, необходимостью обеспечения Советского правительства данными о населении для составления перспективных планов и, во-вторых, тем, что этот год соответствовал рекомендации Статистической комиссии ООН о проведении переписей в год, следующий на юль или близкий к нему. Проведение переписи в январе объясняется тем, что в зимнее время передвижение населения является относительно меньшим, чем в другое время года, что способствует полноте и точности его счета. Кроме того, дата, близкая к началу года, облегчает использование данных переписи для демографических расчетов. Дата переписи (15 января) была выбрана с таким расчетом, чтобы она приходилась на период наименьшей подвижности населения: школьные каникулы к этому времени закончились, а каникулы у студентов вузов и техникумов еще не начались; кроме того, в середине недели (15 января был четверг) подвижность населения меньше, чем в субботу, воскресенье и понедельник.

Критическим моментом переписи было 12 часов ночи с 14 на 15 января. Установление критического момента имеет большое значение для обеспечения точности счета населения. Ведь в нашей стране в январе 1970 г. в среднем каждую минуту рождалось более 8 человек и умирало немногим менее 4 человек. Следовательно, каждую минуту население нашей страны увеличивалось или уменьшалось на 4 человека, т. е. более чем на 6 тыс. человек в сутки. Наряду с естественным движением происходит также и механическое движение населения, в результате чего изменяется распределение населения по районам и населенным пунктам. Поэтому собрание сведений по состоянию на определенный критический момент (день, час) давало возможность избежать как недоучета, так и двойного счета отдельных лиц.

Переписью было охвачено все население СССР. С этой целью переписывалось население не только по месту жительства, но и все граждане, слушающие в ночь с 14 на 15 января в поездах дальнего следования и находящихся в аэропортах, на вокзалах, в больницах, гостиницах, в морских и речных портах. Население, проживающее в труднодоступных и северных районах, было переписано заблаговременно, до прекращения зимой их связи с основными районами страны.

Для обеспечения полноты и правильности счета населения при проведении переписи счетчики в необходимых случаях выдавали *справки о прохождении переписи* и составляли *контрольные бланки*. Справка выдавалась временно проживающим во избежание двойного счета их. С этой же целью справки выдавались также тем лицам, которые предполагали в период с 15 по 29 января выехать в другое место хотя бы на один день.

На избежание недоучета или двойного счета заполнялись контрольные бланки на тех гражданах, которые на критический момент переписи находились в одном месте и должны были пройти там перепись, но справок о прохождении переписи не имели. После окончания переписи контрольные бланки выслаивались в соответствующие пункты для сверки с записями в переписных листах. В результате такой проверки к итогам переписи было добавлено 348 тыс. человек.

После переписи в период с 24 по 29 января включительно инструкторы-контролеры вместе со счетчиками, а в сельских местностях также вместе с уполномоченными сельских Советов с целью проверки правильности счета населения проводили *выборочный контрольный обход* жилых помещений. В городских поселениях контрольный обход проводился во всех счетных участках с охватом в них 50% жилых помещений, а в сельской местности — в 50% счетных участков с охватом в них 100% жилых помещений. Во время контрольных обходов были дополнительно записаны в переписные листы 264 тыс. человек, пропущенные счетчиками. Таким образом, в целом применение всех контрольных мероприятий позволило уточнить данные переписи на 612 тыс. человек (0,25% общей численности населения СССР).

Исчисленные материалы, полученные в результате проведения переписи, были обработаны на ЭВМ «Минск-32» в самый короткий срок.

Бюджетные исследования

Задача бюджетных исследований заключается в изучении размера и структуры доходов и расходов семей различных общественных групп населения. Данные бюджетных исследований используются при планировании, при составлении различных балансов, в том числе при исчислении совокупного общественного продукта и национального дохода, а также для характеристики роста материального благосостояния нации.

Бюджетные исследования представляют собой большую и сложную статистическую работу, которая проводится выборочным порядком (способы и методы отбора семей для бюджетных исследований изложены в гл. VII). Вопросы программы и практики проведения этих исследований рассмотрим на примере исследования бюджетов колхозников.

Оно производится по специально разработанной программе, которая изложена в ряде формуляров (бланков). Программа дает возможность получить характеристику: а) всех денежных и натуральных доходов колхозников по источникам их поступления; б) всех денежных и натуральных расходов по видам затрат; в) личного потребления колхозников; г) источников и размеров приобретений колхозниками промышленных товаров и сельскохозяйственных продуктов. Программой предусматривается также учет затрат труда колхозниками в общественном производстве, в своем хозяйстве и на других работах, вне колхоза.

Обследование проводится специальными работниками статистических органов — статистиками по бюджетам колхозников. Каждый из них проводит обследование бюджетов в 20—23 семьях колхозников путем их опроса и использования документов колхозов, а также записей обследуемых семей в специальных выданных им бланках о денежных поступлениях и расходах.

Статистик по бюджетам два раза в месяц проводит опрос колхозников. При первом опросе (с 16 числа до конца месяца) статистик записывает данные о расходе продуктов на питание, о приходе и расходе денег, о продукции личного подсобного хозяйства за первую половину месяца; при втором (с 1 по 10 число) заполняет все формы бюджетного обследования данными за истекший месяц. В заключение обследования статистик проводит контроль каждого обследования бланка, сопоставляя одноименные показатели в разных формах обследования, проверяя итоги и наличие увязки между различными показателями.

Проверенные бланки представляются в статистическое управление области, края, где производится их разработка.

5. СТРОГАЯ ДОСТОВЕРНОСТЬ И ТЩАТЕЛЬНАЯ ПРОВЕРКА СТАТИСТИЧЕСКИХ ДАННЫХ — ЗАКОН СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ В СССР

Достоверность Отличительной особенностью советской статистики является ее стремление обеспечить строгую достоверность своих данных. К этому призывал советских статистиков В. И. Ленин. Он писал, что обязанностью эконсоветов «является правильное, своевременное представление правдивых отчетов»¹.

Следуя указаниям В. И. Ленина, Коммунистическая партия и Советское правительство требуют того, чтобы в любом звене на-

ционального хозяйства была обеспечена абсолютная достоверность данных о выполнении плана, о развитии всех сторон экономической и культурной жизни страны.

Важную роль в искоренении приписок, очковтирательства — великий глубоко чуждый самой природе нашего общества, имеет указ Президиума Верховного Совета СССР от 24 мая 1961 г. «Об ответственности за приписки и другие искажения отчетности о выполнении плана». Указ устанавливает, что «приписки в государственной отчетности и представление других умышленно искаженных данных о выполнении планов должны рассматриваться как противогосударственные действия, наносящие вред народному хозяйству СССР, и лица, виновные в этом, наказываются лишением свободы на срок до трех лет».

Дальнейшему усилению борьбы за обеспечение неперменной достоверности статистических данных способствует выполнение Постановления ЦК КПСС и Совета Министров СССР «О мерах по усилению контроля фактов обмана государства и по усилению контроля за достоверностью отчетов о выполнении планов и обязательств, принятого в 1967 г.

Ввиду того, что обеспечение абсолютной достоверности отчетности данных является прямой обязанностью работников ЦСУ, строгая ответственность за точность статистических данных лежит на работниках органов государственной статистики.

Достоверность данных советской статистики обеспечивается тем, что она базируется на данных первичного учета. Однако было бы неправильно полагать, что документальность учета чуть ли не автоматически обеспечивает достоверность данных. Чтобы повысить достоверность статистических материалов, нужна продуманная и организованная по правилам статистической науки система контрольных мероприятий. В частности, наряду с проведением работниками органов ЦСУ СССР проверок состояния учета и достоверности данных отчетности на промышленных предприятиях, стройках, в колхозах и совхозах, при контроле отчетности и других материалов наблюдения необходимо применять проверку данных путем балансовой увязки взаимосвязанных показателей в различных отраслях народного хозяйства.

Всякое статистическое наблюдение ставит задачу получения таких данных, которые бы точно отображали действительность. Однако по целому ряду причин степень точности данных наблюдения может быть различной. Неточности, неправильности в статистических данных, полученных при наблюдении, принято называть ошибками наблюдения. Ошибки наблюдения подразделяются на два вида: а) ошибки регистрации и б) ошибки репрезентативности.

Ошибки регистрации могут быть как в сплошном, так и в частичном наблюдении. Они могут быть случайными и систематическими.

¹ Ленин В. И. Полн. собр. соч. т. 43, с. 291.

Случайные ошибки могут быть допущены как опрашиваемыми в их ответах, так и регистраторами при заполнении бланков (на пример, описки). Эти ошибки могут быть направлены как в сторону преуменьшения, так и в сторону преувеличения. При соби- рании сведений о массовых явлениях такие ошибки большую опас- ность не представляют, так как в обобщающих показателях они почти полностью взаимопоглощаются.

Другой характер имеют систематические ошибки, которые обычно направлены в одну определенную сторону. Систематиче- ские ошибки могут быть преднамеренными и непреднамеренными.

Преднамеренные ошибки (вернее, сознательные тенденциозные искажения) получаются в результате того, что опрашиваемый, зная действительное состояние дела, сознательно сообщает регистратору неправильные данные. Например, при переписях садов и ягодников некоторые граждане стремятся преуменьшить количество плодовых деревьев, имеющих в их саду. К преднамеренному искажению данных относятся приписки, очковирательство, допускаемые не- которыми недобросовестными руководителями при составлении отчетности.

Непреднамеренные систематические ошибки допускаются не-умышленно. К ним относятся, например, ошибки, связанные с не- исправностью измерительных приборов, с пропусками в записях и запямятованием, а также ошибки, связанные с тенденцией у ста- риков преувеличивать свой возраст или стремлением некоторых граждан округлять свой возраст до чисел, кратных 5 и особенно 10.

Мерами борьбы с ошибками регистрации являются подбор до- бросовестных и наиболее квалифицированных регистраторов, тща- тельная их подготовка и контроль за их работой, проведение кон- трольных обследований, проверка измерительных приборов, прове- дение разъяснительной работы среди населения и др.

Ошибки репрезентативности возникают только при несплошном наблюдении. Причина их заключается в том, что отобранная и об- следованная часть изучаемой совокупности недостаточно точно отображает состав всей совокупности в целом. Отклонение значе- ния показателя в обследованной части совокупности от размеров его во всей изучаемой совокупности и представляет собой ошибку репрезентативности. Подробнее об этих ошибках при выборочном наблюдении говорится в гл. VII.

Контроль статисти- ческих данных

Полученные в результате статистического наблюдения данные до их обработки под- вергаются самой тщательной проверке. Прежде всего производится так называе- мый внешний контроль. При этом проверяется правильность оформления документов, т. е. наличие и четкость всех необхо- димых записей, предусмотренных инструкцией, а также полнота материала и охвата всех отчетных единиц или единиц наблюдения. Так, например, инструктор-контролер, принимающий за счетчика ма- териалы переписи населения, должен проверить как правильность оформления переписных листов и других документов, так и под-

линту охвата переписью всех домовладений, квартир и хозяйств сельского участка.

После внешнего контроля производится счетный и логический контроль.

Счетный, или арифметический, контроль заключается в счетной проверке всех итоговых показателей, содержащихся в отчетности или в формулярах обследования, и в увязке тех показателей, ко- торые могут быть выведены один из другого. Задачей такого конт- роля является исправление итогов и отдельных числовых пока- зателей.

При контроле результатов статистического наблюдения и осо- бенно отчетности необходимо привлекать материалы из других ис- точников по тем же показателям. Например, данные годовых от- четов можно проверить данными квартальных отчетов, а матери- алы квартальных отчетов — месячными.

В ряде случаев при счетном контроле данных статистическо- го наблюдения применяется метод балансовой увязки показателей (значение на начало отчетного периода плюс поступление минус расход должно равняться наличию на конец отчетного периода). Такой метод контроля применяется, например, при проверке дан- ных о движении поголовья скота в отчетности колхозов и совхозов по животноводству (форма № 24), при проверке отчетности про- мышленных предприятий об остатках, поступлении и расходе сырья в материалах (форма № 1-сн) и т. п.

Логический контроль заключается в проверке ответов на воп- росы программы наблюдения путем их сопоставления между собой или путем сравнения полученных данных с другими источниками на этом же вопросу и т. п. Например, сопоставляя ответы на воп- рос переписного листа о возрасте, семейном состоянии, грамот- ности и другие, можно логическим путем установить ошибку. Если десятилетний мальчик показан женатым или двухлетний ре- беночек показан как грамотный, то ясно, что здесь допущена ошибка в возрасте или в семейном состоянии, или в грамотности. Ответы на другие вопросы переписного листа могут помочь исправить ошибку.

Очень много ошибок выявляется при логическом контроле за- писей в бланках бюджетного обследования. Это достигается сопо- ставлением связанных между собой показателей. Например, если в семье колхозника имеется корова, то в бланке бюджета должны быть обязательно записаны данные о расходе кормов, затратах труда на уход за коровой и т. д. Ошибки в данных по одному рай- ону можно обнаружить при сопоставлении их с данными другого района. При контроле периодической отчетности по промышлен- ности, сельскому хозяйству, строительству и т. д. широко приме- няется сопоставление данных за отчетный период с соответствую- щими данными за предшествующие периоды и с плановыми пока- зателями, а также сопоставление данных из одних форм отчетности с аналогичными или близкими к ним данными из других форм отчетности и других источников (например, данные учета посевных

ОТЧЕТ ПРОМЫШЛЕННОГО ПРЕДПРИЯТИЯ О ВЫПОЛНЕНИИ ПЛАНА ПО ТРУДУ

II КВАРТАЛ 1974 г.

1. Численность и фонд заработной платы персонала

Показатели	Номер строки	Шифр	План			Фактически			Выполнение плана с начала года (стр. 5: стр. 2)	С начала отчетного периода, так к соответствующему периоду прошлого года (стр. 5: стр. 7)
			на год	на квартал с начала года	на отчетные кварталы	с начала года	за соответствующий период с начала квартала прошлого года	за соответствующий период с начала квартала прошлого года		
А	Б	В	1	2	3	4	5	6	7	8
А. Среднесписочная численность работников (включая подростков)										
Персонал предприятия — всего (стр. 2 + стр. 6)	1	101	1110	1117	1123	1098	1212	1135	1135	108,5
Промышленно-производственный персонал — всего	2	110	1070	1077	1083	1171	1175	1097	1098	109,1
в том числе:										
рабочие	3	121	971	978	975	1081	1076	999	999	110,1
ИТР	4	123	×	×	×	×	×	×	×	×
служащие	5	124	×	×	×	×	×	×	×	×
Персонал, занятый в промышленных организациях	6	150	40	40	×	37	37	38	38	92,5
Б. Фонд заработной платы — тыс. руб. с десятичным знаком										
Персонал предприятия — всего (стр. 8 + стр. 13)	7	201	1207,6	620,0	314,0	843,0	670,6	320,3	604,9	108,2
Промышленно-производственный персонал — всего	8	210	1176,0	1604,0	306,0	138,0	1565,0	312,8	—	108,5
в том числе:										
рабочие	9	221	1041,2	536,5	272,3	974,1	582,0	281,3	526,0	108,5
ИТР	10	223	×	×	×	27,9	×	21,5	×	×
служащие	11	224	×	×	×	7,9	×	8,1	×	×
Из стр. 9 в том числе премии рабочим из фонда заработной платы	12	231	×	×	×	40,5	79,2	×	×	×
Персонал, занятый в промышленных организациях	13	250	31,6	16,0	8,0	7,6	15,4	7,5	15,4	96,2
Кроме того, фонд заработной платы работников несписочного состава	14	280	0,6	0,3	0,2	—	0,3	—	0,1	100,0
В. Премии, единовременные поощрения и вознаграждения, начисленные из фонда материального поощрения (без премий рабочим по фонду заработной платы и единовременной помощи) — тыс. руб. с десятичным знаком										
Всему персоналу	15	301	—	—	—	—	—	—	—	—
Промышленно-производственному персоналу	16	310	—	—	—	—	—	—	—	—

А	Б	В	1	2	3
в том числе:					
рабочим	17	321	—	—	—
ИТР	18	333	×	×	×
служащим	19	324	×	×	×
Г. Фонд заработной платы, премии и единовременные поощрения — тыс. руб. с десятичным знаком					
Промышленно-производственному персоналу (стр. 8+стр. 16)	20	370	1176,0	604,0	306,0
в том числе:					
рабочим (стр. 9+стр. 17)	21	381	1041,2	536,5	272,3
ИТР (стр. 10+стр. 18)	22	383	×	×	×
служащим (стр. 11+стр. 19)	23	384	×	×	×

4	5	6	7	8	9
в том числе:					
рабочим	—	—	—	—	—
ИТР	×	×	×	×	×
служащим	×	×	×	×	×
Г. Фонд заработной платы, премии и единовременные поощрения — тыс. руб. с десятичным знаком					
Промышленно-производственному персоналу (стр. 8+стр. 16)	335,0	655,2	312,8	589,5	108,5
в том числе:					
рабочим (стр. 9+стр. 17)	299,1	582,0	281,3	526,0	108,5
ИТР (стр. 10+стр. 18)	27,0	×	21,5	×	×
служащим (стр. 11+стр. 19)	7,0	×	8,1	×	×

площадей можно сопоставить с данными земельного баланса и т. п.).

Последовательность осуществления контроля и методы выявления ошибок рассмотрим на примере. Допустим, что в статистическое управление 8 июля 1974 г. одной из швейных фабрик был представлен отчет по форме № 2-т о выполнении плана по труду за II квартал 1974 г. Первый раздел этого отчета содержал следующие данные (см. табл. 3.1).

Приступая к проверке отчета, прежде всего производем внешний контроль, т. е. проверяем полноту заполнения всех строк и граф, которые по инструкции обязательно должны содержать те или иные данные. Устанавливаем, что в проверяемом отчете по неизвестной причине оказались незаполненными: по стр. 6 гр. 3 и по стр. 8 гр. 7. Руководствуясь указанными инструкцией по составлению данной статистической отчетности, подсказками, содержащимися в самой форме отчета, а также используя имеющийся в статуправлении отчет данного предприятия за прошлый квартал, можно предварительно восполнить проверяемый отчет недостающими сведениями без запроса предприятия. Так, исходя из того, что по каждой графе данные по стр. 1 должны быть равны сумме данных по стр. 2 и 6, устанавливаем, что по стр. 6 гр. 3 должно быть записано 40 человек (1123—1083). Аналогично восполняем данные стр. 8 гр. 7. Найденные недостающие сведения записываются в проверяемый отчет лишь после получения ответа предприятия на соответствующий запрос (по почте или по телефону с последующим письменным подтверждением).

После проверки полноты заполнения отчета приступаем к счетному и логическому контролю содержащихся в нем данных.

При счетной проверке правильности итоговых данных по стр. 1 путем суммирования данных стр. 2 и 6 обнаруживаем отсутствие необходимого равенства в гр. 4. Чтобы установить, где именно допущена ошибка — в слагаемых или в итоге (сумма), сопоставим фактическую численность всего персонала предприятия за отчетный квартал — 208 человек (стр. 1 гр. 4) с соответствующими данными по плану — 123 человека (стр. 1 гр. 3) и с фактическими данными за отчет за предыдущий (первый) квартал — 1203 человека. Это сопоставление приводит к выводу, что именно итог является ошибочным, и вместо 208 человек по стр. 1 гр. 4 должно быть записано 1208 человек, т. е. 1117+37 (сумма стр. 2 и 6).

Далее проверяем данные стр. 7, 8 и 13, для которых должно иметь место равенство: стр. 7 = стр. 8 + стр. 13. В проверяемом отчете такого равенства

не получается в гр. 2 и 5. Причем в каждом случае ошибочными могут оказаться данные как стр. 7, так и стр. 8 или 13. Чтобы выяснить, в какой строке гр. 3 записаны ошибочные данные, сопоставим плановый фонд заработной платы за квартал с начала года, т. е. на первое полугодие (гр. 2), с плановым фондом на июль 1974 г. (гр. 1) по стр. 7, 8 и 13. Так как фонд заработной платы на полугодие не может быть больше годового фонда, как это получается по стр. 8, то очевидно, что по этой строке в гр. 2 приведены неправильные данные (по стр. 7 и 13 такого противоречия нет). Зная, что сумма заработной платы, показанная по стр. 7, должна быть равна сумме данных стр. 8 и 13, находим действительную величину показателя по стр. 8 гр. 2: стр. 8 — стр. 7 — стр. 13 = 160,0—16,0 = 604,0 (тыс. руб.).

Чтобы убедиться в правильности полученного результата, рассмотрим показатели стр. 15, 16 и 17. Ясно, что при отсутствии записей по стр. 16 (премия, единовременные поощрения и вознаграждения, начисленные из фонда материального поощрения промышленно-производственному персоналу) суммы, показанные по стр. 8 и 20, должны быть равны между собой. Следовательно, найденная нами сумма (604 тыс. руб.), совпадающая с данными стр. 20 гр. 2, действительно является правильной величиной фонда заработной платы промышленно-производственного персонала по плану на первое полугодие (стр. 8 гр. 2).

По гр. 5, где также не соблюдается равенство стр. 7 суммы стр. 8 и 13, сразу выявляет сомнение сумма, показанная по стр. 8, которая более чем в три раза превышает гр. 4 той же строки и даже больше запланированной суммы заработной платы промышленно-производственному персоналу на год (гр. 1 стр. 4). Кроме того, данные стр. 8 гр. 5 не совпадают с данными стр. 20 той же строки. Действительная сумма, которая должна быть записана по стр. 8 в гр. 5, является так же, как это было сделано по гр. 2.

На гр. 4 стр. 9 устанавливаем ошибку в том, что часть оказалась больше плана: фонд заработной платы рабочих, являющихся частью промышленно-производственного персонала, оказался больше фонда заработной платы как промышленно-производственного, так и всего персонала предприятия. Действительную сумму фонда заработной платы по стр. 9 устанавливаем исходя из того, что стр. 21 = стр. 9 + стр. 17.

Важнейшие исправления данных в проверяемый отчет вносятся лишь после получения ответа от предприятия на запрос по поводу допущенных ошибок. В прилагаемом в разработку отчете будут такие исправления и дополни-

ния в табл. 3.1: а) стр. 1 гр. 4—1208; б) стр. 6 гр. 3—40; в) стр. 8 гр. 2—604,0; г) стр. 8 гр. 5—655,2; д) стр. 8 гр. 7—589,5; е) стр. 9 гр. 4—298,1.

С целью повышения качества отчетности, представляемой предприятиями, организациями и учреждениями, статистические органы по окончании проверки и разработки той или другой формы отчетности направляют предприятиям письма, в которых наряду с перечнем выявленных ошибок указывают причины, их породившие, и излагают пути избежания их в дальнейшем. С работниками предприятий и организаций, отвечающими за сроки представления и качество составления статистической отчетности, систематически проводятся совещания, на которых обсуждаются результаты проверок достоверности отчетных данных и пути улучшения работы по учету и отчетности.

Глава IV

СВОДКА И ГРУППИРОВКА МАТЕРИАЛОВ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ

Выводы

1 СОДЕРЖАНИЕ И ЗАДАЧИ СТАТИСТИЧЕСКОЙ СВОДКИ

Сводка — второй
этап
статистического
исследования

После того как первичный статистический материал собран и проконтролирован, можно перейти ко второму этапу статистического исследования — сводке материалов статистического наблюдения.

Статистическое наблюдение дает материал, характеризующий отдельные единицы объекта исследования. Задача сводки — подытожить, систематизировать и обобщить результаты наблюдения так, чтобы стало возможным выявить характерные черты статистической совокупности в целом и обнаружить закономерности изучаемых явлений и процессов.

Например, в результате переписи населения собирают определенные сведения о каждом отдельном жителе страны: пол, возраст, состоит ли в браке, национальность, родной язык, образование и т. д. Во время переписи населения СССР в 1970 г. были собраны такие сведения о каждом из 241 720 тыс. человек. Но эти данные в своем первоначальном виде представляют лишь исходный материал для получения сводных, обобщающих характеристик населения страны как статистической совокупности. Чтобы судить не об отдельных жителях страны, а дать характеристику населения страны в целом, т. е. чтобы установить состав населения по основным группам, занятиям, полу, возрасту, национальности, образованию и т. д., а именно в этом и заключается задача статистики, нужно собранный материал об отдельных жителях страны систематизировать, обобщить, привести в определенный порядок. То же самое обстоит дело и в любом ином статистическом исследовании. Получение сведений о каждой единице совокупности служит средством для достижения более широкой цели — получению обобщающей характеристики статистической совокупности в целом.

Следовательно, действия по упорядочению и обработке первичного статистического материала в целях выявления типических черт и закономерностей изучаемых явлений и процессов представляют содержание сводки материалов статистического наблюдения.

В статистической литературе иногда различают два понятия сводки: сводку в широком смысле слова и сводку в узком смысле слова. Под *сводкой в широком смысле слова* понимают большой круг операций, связанных с получением типичных характеристик массовых общественных явлений и процессов. Она включает в себя распределение единиц изучаемого объекта на группы, характеристику этих групп с помощью системы показателей, а также изложение результатов в виде статистических рядов, таблиц и графиков. Под *сводкой в узком смысле слова* понимается подсчет групповых и общих итогов.

Научно обоснованная организация сводки

Только при научно обоснованной организации статистической сводки можно получить правильную, соответствующую действительному положению вещей характеристику изучаемых общественных явлений. Наоборот, при неумелой, непродуманной сводке первичный материал, каким бы богатым он ни был, теряется, обесценивается и, говоря словами В. И. Ленина, «экономист получает в свое распоряжение рутинные, бессмысленные столбцы цифр, статистическую «игру в цифирки» вместо осмысленной статистической обработки материала»¹.

Вот почему вопросы научного обоснования и правильной организации сводки относятся к числу основных вопросов статистической методологии, с которыми экономист сталкивается при любом исследовании общественных явлений и процессов.

Давая оценку деятельности русской дореволюционной земской статистики, В. И. Ленин отмечал, что земскими статистиками были собраны превосходные по своей детальности, по тщательности собирания и проверки массовые данные. Теперь, указывая В. И. Ленин, задача заключается в том, чтобы эти материалы были обработаны надлежащим образом, т. е. так, «чтобы *получалась ответ, точный, объективный, основанный на учете массовых данных, ответ на все вопросы, указанные или намеченные* более чем полувековым анализом пореформенной экономики России...»². Вместе с тем В. И. Ленин неоднократно критиковал дореволюционную земскую статистику за плохую, ненаучную организацию сводки, которая губила ценный первичный материал.

С критикой ненаучной сводки материалов статистического наблюдения В. И. Ленин выступал и в адрес американской правительственной статистики. Говоря о разработке материалов американской сельскохозяйственной переписи 1910 г., В. И. Ленин характеризовал ее как нагляднейший образец того, как великолепный по богатству и полноте материал обесценен, испорчен рутинной и научным невежеством тех, кто его обрабатывал³.

В. И. Ленин требовал, чтобы сводка была осмысленной. Критикуя работу Козьминых-Ланина «Рабочий день и рабочий год в Московской губернии», изданной в 1912 г., Ленин отмечал, что «изложение у автора не дает общей картины, что получилось «увлечение статистикой ради статистики, своего рода игра в цифири — в ущерб ясности картины и пригодности материала для изучения»⁴.

Вопросы правильной организации сводки не потеряли своей значимости сегодня. Исключительно важное значение вопроса о научной организации сводки для советской статистики вытекает из того, что она в результате наблюдения получает богатейшие данные о состоянии и развитии советской экономики и культуры, о ходе выполнения народнохозяйственных планов и т. д. Поэтому только научно обоснованная их сводка с широким использованием различных группировок позволит при последующем анализе выявить те или иные закономерности, обнаружить взаимосвязи, измерить влияние различных факторов на достигнутые результаты и учесть все это в практической работе, в том числе при составлении текущих и перспективных планов.

Успех статистической сводки во многом зависят ее программа и план. В самом общем виде программа статистической сводки содержит перечень групп, на которые нужно рассчитать совокупность, а также перечень показателей, используемых для характеристики изучаемой совокупности в целом и отдельных ее частей.

В зависимости от задач, поставленных перед данным конкретным исследованием, программа устанавливает группировочные признаки для образования однородных в том или ином отношении групп, число групп и их границы, макеты разработанных таблиц, содержащие как наименования групп, так и необходимые для их характеристики показатели. Программа сводки предусматривает также тот или иной территориальный ее разрез, т. е. выделение территориальных единиц (районов, областей и т. д.).

Общественные явления весьма многогранны. Поэтому программа должна быть составлена так, чтобы в результате сводки получить материал, характеризующий изучаемое явление с различных сторон. Например, в результате сводки данных переписи населения можно получить материал о половом и возрастном составе населения, о классовом его составе, культурном уровне и т. д. Для этого при составлении программы сводки необходимо сотрудничество ряда учреждений, широкого круга специалистов. В составлении программы сводки обязательно должны участвовать и те специалисты, которые будут использовать материал сводки.

Наряду с программой успех сводки во многом зависит от плана ее проведения. Как и всякая большая и сложная работа, статистическая сводка должна осуществляться по заранее составленному

¹ Ленин В. И. Полн. собр. соч., т. 27, с. 182.

² Там же, т. 24, с. 276.

³ См.: Там же, т. 27, с. 182.

⁴ Ленин В. И. Полн. собр. соч., т. 22, с. 31.

плану. В данном случае это особенно важно потому, что в статистической сводке, как правило, участвует большое число лиц. План должен содержать указания о последовательности и сроках выполнения отдельных частей сводки, ее исполнителях и о порядке изложения ее результатов.

2. СТАТИСТИЧЕСКИЕ ГРУППИРОВКИ

Группировки — научная основа сводки

В сводке статистического материала вопрос о группировке представляет самое важное звено, так как простой подсчет итогов без распределения единиц совокупности на группы по тем или иным признакам не дает полной характеристики объекта изучения. Например, если при сводке материалов переписи населения ограничиться только подсчетом числа жителей, то будут получены лишь сведения об общей численности населения того или иного района или страны в целом. Этих сведений, конечно, недостаточно для анализа и выявления закономерностей. Чтобы охарактеризовать население подробно и глубоко, нужно знать, как оно распределяется по полу, возрасту, общественным группам, занятиям и т. д. Сказанное относится к любой статистической сводке.

Поэтому правильная организация статистической сводки предполагает предварительное, т. е. до производства подсчетов, выделение из общей совокупности характерных ее частей, предварительную группировку единиц совокупности. Распределение единиц изучаемого объекта на однородные группы по существенным для них признакам называется статистической группировкой.

Метод статистических группировок, выступая как первый шаг статистической сводки, позволяет разработать первичный статистический материал так, чтобы получили четкое выражение все существенные черты и особенности изучаемых явлений. Этим определяется роль группировок как научной основы сводки.

В научных исследованиях и в практической экономической работе группировки применяются давно. В подтверждение можно сослаться на таких известных прогрессивных общественных деятелей и ученых, как А. И. Радищев, Н. П. Огарев, Д. П. Журавский и ряд других, которые обращались к методу статистических группировок при изучении общественных явлений. Так, Д. П. Журавский отмечал, что статистическое исследование носит характер счета по группам, категориям. Известно, что он и статистику определял как науку о «категорическом исчислении».

Исключительно важное значение для статистической науки имеет ленинская трактовка вопроса о группировках. В противоположность буржуазным статистикам, которые видели в группировках простой технический прием расчленения целого на части, В. И. Ленин придавал большое значение статистическим группировкам. Ленин показал, что один и тот же исходный материал приводит к диаметрально противоположным выводам при различных

приемах группировки, что, следовательно, эти приемы должны определяться не субъективными построениями статистика, а объективными исследуемыми явлениями.

Научные группировки должны опираться на всесторонний анализ изучаемых явлений с позиций экономической теории. Раскрытие закономерностей экономического развития могут дать только те группировки, которые исходят из законов, устанавливающих закономерностей политической экономии. В. И. Ленин подчеркивал, что вопрос о группировках статистического материала «... вовсе не является таким узкотехническим, узкоспециальным вопросом, каков он может показаться на первый взгляд»¹, что о приемах статистических группировок «... вправе говорить и даже обязаны говорить вовсе не одни земские статистики, но и все экономисты»².

Метод группировок применяется для решения различных задач, возникающих в ходе научного статистического исследования.

Важнейшими из них являются: 1) выделение социально-экономических типов; 2) определение структуры однородных совокупностей и 3) выявление связей и зависимостей между отдельными признаками общественных явлений. Соответственно этому различают три вида группировок: 1) типологические, 2) структурные и 3) аналитические.

Группировка единиц совокупности в зависимости от задач исследования может производиться по одному или нескольким признакам. В том случае, когда группы образуются по одному признаку, группировка называется простой. Примером простой группировки может служить распределение населения СССР по социальным классам:

Таблица 4.1

ВЫСШЕГО СОСТАВА НАСЕЛЕНИЯ СССР (процентов) -

	1913 г.	1929 г.	1959 г.	1972 г.
Все население (включая неработающих членов семей)	100	100	100	100
в том числе:				
рабочие и служащие	17,0	50,2	68,3	82,2
из них рабочие	14,6	33,5	49,5	60,6
крестьяне, крестьянство и кооперированные кустари	—	47,2	31,4	17,8
крестьяне-единоличники и некооперированные кустари	66,7	2,6	0,3	0,0
предприниматели, помещики, торговцы и другие	16,3	—	—	—

¹ Ленин В. И. Полн. собр. соч., т. 27, с. 182.
² Там же, т. 3, с. 632.

Если же группы образуются на основе двух или большего числа признаков, взятых в комбинации друг с другом, т. е. группы, образованные по одному признаку, делятся затем на подгруппы по второму и т. д. признакам, то такие группировки называются **комбинационными**. Приведем пример комбинационной группировки.

Таблица 4.2
ГРУППИРОВКА КОЛХОЗОВ ОБЛАСТИ ПО ВАЛОВОМУ ДОХОДУ НА ЧЕЛОВЕКО-ДЕНЬ
И ВАЛОВОЙ ПРОДУКЦИИ НА ГЕКТАР СЕЛЬСКОХОЗЯЙСТВЕННЫХ УГОДИЙ
ЗА ОТЧЕТНЫЙ ГОД
(в процентах к итогу)

Группы колхозов по размеру валового дохода на 1 чел.-день (руб.)	Распределение колхозов по валовой продукции на 1 га сель- скохозяйственных угодий (руб.)			
	менее 300	300—500	500 и выше	и то г о
I. Менее 5,00	8	18	4	30
II. От 5,00 до 7,50	12	21	22	55
III. 7,50 и более	—	6	9	15
Всего	20	45	35	100

Комбинационные группировки особенно необходимы в тех случаях, когда речь идет об изучении сложного экономического явления или процесса, например, при изучении технического прогресса в промышленности, путей развития сельского хозяйства и т. д. В этих случаях одновременное использование нескольких группировочных признаков, взятых в сочетании друг с другом, позволяет более полно и всесторонне охарактеризовать изучаемое явление, чем на основе изолированной группировки по ряду группировочных признаков.

Типологические группировки

Из ленинской теории группировок вытекает, что обобщающие статистические характеристики могут применяться лишь к качественно однородным совокупностям. Поэтому первая задача метода группировок заключается в том, чтобы на основе экономической теории распределить массу изучаемых единиц на однородные в социально-экономическом отношении группы, выделить социально-экономические типы явлений и процессов. Группировки, приводящие к выделению социально-экономических типов, называются **типологическими**. Они широко применяются при изучении различных общественных явлений. Здесь в первую очередь назовем группировку населения по социальному признаку — классам (см. табл. 4.1). Из таблицы видно, что в настоящее время в СССР практически 100% населения связано с социалистическим хозяйством.

На современном этапе строительства коммунизма важное значение имеет группировка по общественным формам собственности.

Советская статистика выделяет государственную социалистическую собственность, являющуюся всенародным достоянием, и колхозную социалистическую собственность, являющуюся собственностью отдельных коллективов трудящихся, членов колхоза. Отдельно показывается мелкая частная собственность крестьян-единоличников, основанная на личном труде. Наконец, для самостоятельного изучения выделяется в особую группу личная собственность колхозников, исключаящая эксплуатацию чужого труда и носящая подсобный характер.

На основе типологических группировок дается сравнительная характеристика эффективности деятельности различных социально-экономических типов хозяйств. Например, при изучении производительности труда в сельском хозяйстве можно установить, что производительность труда в совхозах выше, чем в колхозах.

Таблица 4.3

ПРЯМЫЕ ЗАТРАТЫ ТРУДА НА ПРОИЗВОДСТВО 1 ц ПРОДУКЦИИ
В КОЛХОЗАХ И СОВХОЗАХ МИНИСТЕРСТВА
СЕЛЬСКОГО ХОЗЯЙСТВА СССР В 1973 г.
(в человеко-часах)

	Зерно (без кукурузы)	Хлопок-сырец	Молоко	Шерсть	Привес	
					крупного рогатого скота	свиней
Колхозы	1,5	36	11	279	61	45
Совхозы	1,1	27	9	200	43	32

Для научно обоснованного построения типологических группировок важное значение имеет правильный выбор группировочных признаков, т. е. признаков, на основе которых производится распределение единиц объекта на однородные группы. От правильного, научно обоснованного выбора этих признаков зависят выводы, получаемые в результате статистической сводки. Поэтому из множества признаков, которыми обладает объект, нужно выбирать самые существенные для целей данного исследования.

Например, при исследовании внутрипроизводственных резервов производительности необходимо выделить передовые предприятия и сравнить показатели их работы с соответствующими показателями для всей промышленности. Выделить типы передовых предприятий можно с помощью таких признаков: уровень рентабельности, систематический рост производительности труда и снижение себестоимости продукции, применение передовой технологии, высокое качество продукции и др. Такие же признаки, как число работающих, объем продукции и т. д. являются в данном случае, т. е. для выделения передовых предприятий, второстепенными, несущественными. Однако при изучении связи между размером предприятий (малые, средние, крупные) и производительностью труда или се-

бестоймостью единицы продукции число работающих и объем продукции становятся существенными признаками.

Отбор признаков, которые кладутся в основу типологической группировки, должен основываться на предварительном глубоком анализе сущности изучаемого явления или процесса. Без этого выбор группировочных признаков превращается в формальную техническую операцию, лишая группировку научной значимости. Необходимо также выделять группировочные признаки в зависимости от конкретных условий. Например, при группировке промышленных предприятий различных отраслей промышленности на крупные, средние и мелкие необходимо брать различные группировочные признаки. Если для химической и металлургической промышленности в основу группировки может быть положен признак мощности оборудования, который определяет объем выпускаемой продукции, то для машиностроительной промышленности группировочным признаком должна служить численность рабочих, которая находится в большей связи с объемом продукции, чем мощность оборудования.

Наконец, составление типологических группировок может потребовать одновременного использования не одного, а нескольких группировочных признаков. Так, В. И. Ленин считал, что для всестороннего изучения крестьянских хозяйств необходима «... во-первых, группировка хозяйства не по одному признаку (величина площади), а по нескольким признакам (количество машин, скота, площади под специальными культурами и пр.), а во-вторых, комбинирование различных группировок, т. е. разделение каждой группы, напр., по величине площади, на подгруппы по количеству скота и т. д.»¹.

Классические образцы научно обоснованного выбора группировочных признаков содержат работы В. И. Ленина. В его трудах выбор группировочных признаков выступает как неотъемлемая составная часть политико-экономического исследования исторически конкретных общественных явлений. Так, при изучении состава населения России по данным переписи 1897 г. В. И. Ленин отвергает формальную группировку населения на 65 групп, приведенную в официальном издании, и группировку население по классовому положению (пролетарии и полупролетарии, беднейшие мелкие хозяйства, крупная буржуазия, помещики и т. д.), а также по отраслям хозяйства (сельскохозяйственное, торговое-промышленное, производственное).

Анализируя развитие капитализма в промышленности России, В. И. Ленин применил научную группировку предприятий на фабричные, ремесленные заведения и кустарные промыслы. Это позволило ему доказать, вопреки утверждениям народников, что число фабрик, объем их продукции, количество фабричных рабочих быстро возрастают. Критикуя ненаучные группировки крестьянских промыслов, составившиеся земской статистикой, В. И. Ленин отмечал, что до тех пор, пока «промыслы» не будут распределены по

их экономическим типам, пока не будут отделяться хозяйства от наемных рабочих, экономическая статистика не может быть признана удовлетворительной.

В. И. Ленин считал, что экономическая статистика должна положить в основу группировки крестьянских хозяйств размеры и типы хозяйств сообразно с местными условиями и формами земледелия. Он критиковал земских статистиков за то, что «вместо изучения типов крестьянского хозяйства (поденщик, средний крестьянин, предприниматель) они изучают, как любители, бесконечные таблицы цифр, точно задавшись целью удивить мир своим арифметическим усердием»².

Изучая процесс развития капитализма в сельском хозяйстве, В. И. Ленин, опираясь на экономическую теорию, делит крестьянские хозяйства на три группы: 1) хозяйства, основанные на эксплуатации, т. е. хозяйства капиталистического типа, 2) хозяйства арендаторские и 3) хозяйства бедняцкие, наиболее жестоко эксплуатируемые. Положив в основу деления признак экономической мощи, т. е. социальной принадлежности, В. И. Ленин убедительно показал, как далеко зашел процесс расслоения крестьянского хозяйства и развития капитализма в сельском хозяйстве России.

С этих позиций В. И. Ленин подверг резкой критике ненаучные группировки крестьянских хозяйств, в основу которых были положены такие несущественные признаки, как число членов семьи или размер земельного надела. Действительно, по числу членов семьи нельзя судить о том, относится ли данное хозяйство к зажиточным или бедняцким. Группировка по наделу также не давала возможности судить о социальной принадлежности хозяйства, так как сама она определялась числом членов семьи. Следовательно, такие группировки затуманивали классовое расслоение деревни.

Критикуя группировки по размеру надела, к которым так часто прибегали народники, чтобы доказать, будто капитализм в России — случайное явление, В. И. Ленин отмечал, что, пользуясь этой группировкой, мы складываем вместе бедняков и богачей, мелких пролетариев с представителями сельской буржуазии.

Острой критике В. И. Ленин подвергал также группировки западноевропейской и американской правительственной статистики за общность с русской земской статистикой недостаток — отсутствие экономического обоснования выбора группировочных признаков, сведение статистических группировок к чисто техническому приему деления совокупности единиц на некоторое число частей.

Исход за выбором группировочных признаков важнейшее значение в методологии построения типологических группировок имеет выявление качественно однородных групп, установление их количества и границ. Эти вопросы должны решаться на основе экономической теории, в соответствии с конкретными целями исследования и особенностями изучаемого объекта.

¹ Ленин В. И. Полн. собр. соч., т. 5, с. 213.

² Ленин В. И. Полн. собр. соч., т. 3, с. 140.

В случае группировки по атрибутивному признаку число групп должно соответствовать указаниям экономической теории и числу разновидностей признака. Так, при группировке населения СССР по социальному положению образуются такие группы, как рабочие, служащие, колхозные крестьяне, крестьяне-единоличники и др. Сельскохозяйственные предприятия нашей страны делятся на колхозы, совхозы и другие государственные хозяйства. В ряде случаев выделяются личные подсобные хозяйства колхозников, рабочих, служащих и других групп населения.

В случае же группировки по количественным признакам необходимо на основе экономической теории и в соответствии с задачами исследования установить количество групп и определить те значения варьирующего признака, которые отделяют один тип от другого и являются границами типических групп. Иными словами, нужно найти такие значения уровней варьирующего признака, переход через которые означает переход к иному социально-экономическому типу.

Так, например, при анализе работы промышленных предприятий с целью вскрытия внутрипроизводственных резервов возможна типологическая группировка на передовые и отстающие предприятия. В качестве группировочного признака здесь может выступить такой существенный признак, как процент выполнения плана по выпуску продукции. Уровнем варьирующего признака, который разделяет эти два типа предприятий, будет 100%. Предприятия, которые не достигли этого рубежа, которые выполнили план по выпуску продукции меньше чем на 100%, т. е. не выполнили планового задания, составляют один тип предприятия, предприятия же, выполнившие план по выпуску продукции на 100% и более — другой тип предприятия.

Экономической теорией доказано, что так называемые качественные показатели (производительность труда работающих, себестоимость единицы продукции, рентабельность и др.) при прочих равных условиях лучше на тех предприятиях, которые выполняют и перевыполняют плановое задание по выпуску продукции. Следовательно, количественное изменение процента выполнения плана по выпуску продукции с 99 до 100% влечет за собой существенное изменение основных показателей работы предприятия.

Большое значение правильному выбору уровней варьирующего признака, отделяющих один тип от другого, придавал В. И. Ленин. При группировке крестьянских хозяйств по экономической состоятельности на основе такого признака, как наличие рабочего скота, Ленин образовал три типа домохозяйств: 1) низшую группу — без рабочего скота и с 1 головой рабочего скота, 2) среднюю группу — с 2—4 головками скота и 3) высшую группу — с 5 и более головками скота¹. В условиях обработки земли с помощью живой тягловой силы переход от одной такой группы к другой означал не только

количественное изменение, но и качественное, существенное изменение в ведении хозяйства: переход от низшей, бедняцкой группы крестьянства к промежуточной, середняцкой и далее — к зажиточному крестьянству, ведущему крупное хозяйство кулацкого, капиталистического типа.

При типологической группировке на основе количественных признаков интервалы, относящиеся к тому или иному типу, должны учитывать особенности объектов, по которым производится группировка. Например, группировка промышленных предприятий на мелкие, средние и крупные по такому признаку, как число работающих, будет иметь различные интервалы для полиграфической промышленности и, допустим, машиностроения. Если машиностроительное предприятие с числом работающих 500 человек является мелким, то полиграфическое предприятие с таким же числом работающих не является уже мелким предприятием. Следовательно, возникает необходимость специализации интервалов на основе учета особенностей каждой из отраслей промышленности².

Замечательный образец применения специализированных интервалов содержит работа В. И. Ленина «Развитие капитализма в России». Анализируя материалы подворных переписей кустарей Московской губернии, содержащие данные по 33 различным промыслам, В. И. Ленин устанавливает существенные в них различия. Учитывая эти различия, В. И. Ленин при группировке кустарных заведений на разряды по такому признаку, как число рабочих, применяет специализированные интервалы. В очень мелких промыслах к низшему разряду он относит заведения с 1 рабочим, к среднему — с 2, к высшему — с 3 и более, а в более крупных промыслах к низшему — заведения с 1—5 рабочими, к среднему с 6—10 и т. д. «Без применения различных приемов группировки, — писал В. И. Ленин, — мы не могли бы представить по каждому промыслу данные о заведениях различной величины»³.

Второй важнейшей задачей статистических группировок является исследование структуры типически однородных групп, изучение вариаций признаков внутри одной типической совокупности. Группировки, характеризующие распределение единиц одной типической совокупности по каким-либо признакам, называются структурными. К ним относятся группировки рабочих по возрасту, стажу работы, заработной плате и т. д., группировки животных по поголовью скота, площади пашни и т. д.

Взятые за ряд периодов (или моментов) времени структурные группировки показывают изменения структуры изучаемых явлений.

Структурная группировка возможна как по одному, так и по нескольким группировочным признакам в их сочетании, комбинации. Например, для анализа состава рабочих предприятия целесообразно производить группировку не только по отдельно взя-

¹ См.: Ленин В. И. Полн. собр. соч., т. 3, с. 84.

² Ленин В. И. Полн. собр. соч., т. 3, с. 342—343.

тым признакам (возраст, стаж работы, квалификация и т. д.), но и в сочетании таких признаков (по возрасту и квалификации, квалификации и заработной плате и т. д.).

Построение структурных группировок, так же как и типологических, возможно как на основе атрибутивных, так и на основе количественных признаков. При группировке по атрибутивному признаку группы отличаются друг от друга не по величине, а по характеру признака. Число групп, на которые делится изучаемая совокупность, часто определяется при этом числом разновидностей атрибутивного признака. Так, группировка рабочих по полу дает две группы, а группировка рабочих по профессиям допускает образование столько же групп, сколько различных профессий имеют рабочие данного предприятия или отрасли.

К числу атрибутивных группировок относятся и особого рода сложные группировки, основанные на статистических классификациях — устойчивых перечнях однотипных групп и подгрупп, образованных на основе тех или иных признаков. Например, народное хозяйство подразделяется на отрасли: промышленность, сельское хозяйство и т. д., а каждая крупная отрасль делится в свою очередь на более узкие отрасли; работающие в промышленности по выполняемым функциям подразделяются на категории: рабочие, инженерно-технические работники, служащие и т. д.

Такого рода классификации обычно разрабатываются централизованно, сопровождаются подробными инструкциями и применяются в течение ряда лет. Классификации, применяемые советской статистикой, отличаются научной обоснованностью и народнохозяйственным принципом построения, который позволяет сопоставлять статистические материалы различных отраслей народного хозяйства и сводить их в едином балансе.

При составлении структурных группировок на основе количественных признаков необходимо определять оптимальное число групп и ширину интервалов.

Вопрос о числе групп решается в этом случае применительно к конкретным задачам исследования. При группировке рабочих-сдельщиков по размеру выполнения ими норм выработки могут быть поставлены разные задачи, что вызовет различное решение вопроса об оптимальном числе групп. Так, может быть поставлена задача выявления невыполняющих норм выработки и выполняющих и перевыполняющих их. Тогда следует образовать две группы. Может быть поставлена задача выявления рабочих, не выполняющих нормы выработки, значительно перевыполняющих нормы выработки и, наконец, находящихся посередине между этими двумя полярными группами. Тогда следует образовать три группы. Наконец, может быть поставлена задача изучения характера вариации процента выполнения норм и характера распределения рабочих по этому признаку. В этом случае исходя из объема изучаемой совокупности следует установить такое количество групп, чтобы достаточно четко выявился характер распределения.

В математической статистике рекомендуется следующая формула оптимального числа групп: $K = 1 + 3,2 \log n$, где K — оптимальное число групп и n — число единиц совокупности.

При определении числа групп необходимо стремиться к тому, чтобы не исчезли особенности изучаемого явления. Например, при изучении зависимости производительности труда от размера машиностроительных предприятий по числу работающих необходимо, чтобы число групп было не слишком большим и не слишком малым и чтобы при этом в каждую группу вошло достаточно большое число предприятий.

Если в данном случае образовать несколько десятков групп (до 100 человек, 101—200 человек, 201—300 человек и т. д.), то в них растворятся особенности, выражающие влияние размеров предприятий на производительность труда. С другой стороны, при малом числе групп в одну и ту же группу попадут предприятия с различным выражением изучаемой особенности. Так, если при исследуемой группировке образовать всего две группы (с числом работников до 5000 и затем свыше 5000 человек), то в первую группу попадут не только мелкие и средние, но частично и крупные предприятия, что также сделает невозможным выявить влияние размеров предприятий на производительность труда.

Что касается требования, чтобы в каждую группу попало достаточно большое число единиц, то оно вытекает из закона больших чисел. Вместе с тем конкретное определение достаточной численности групп нужно соотносить с сущностью изучаемого объекта и задачами исследования. В отдельных случаях интерес представляют и малочисленные группы, в частности, когда идет речь об образовании нового, которое только зарождается и не имеет еще массового характера. Малочисленные группы представляют интерес и в ряде других случаев. Например, в США 1% крупных собственников владеет примерно 60% всего национального богатства. Хотя эта группа малочисленна, ее необходимо выделить в самостоятельную при изучении характерных черт капитализма. Следовательно, при решении вопроса о численности единиц в группах нужно руководствоваться не формальными соображениями, а знанием сущности изучаемого вопроса, требованиями марксистско-ленинской теории.

При группировке по количественным признакам возникает также вопрос о размере интервала. Под величиной интервала понимается разность между наибольшим и наименьшим значениями признака в каждой группе. Здесь также нужно руководствоваться сущностью изучаемого явления, чтобы получить группы, существующие в действительности, четко отличающиеся одна от другой. Например, распределяя рабочих промышленности по возрасту, выделяют группы до 18 лет (с сокращенным рабочим днем), 18, 19 и 20 лет (чтобы видеть степень вовлечения молодежи в промышленность), затем образуют более крупные интервалы — 21—24 года, 25—34 года и т. д. и заканчивают группой 60 лет и старше (пенсионный возраст).

Интервалы групп могут быть *равные и неравные*. При исследовании экономических явлений обычно применяются неравные, чаще всего прогрессивно увеличивающиеся интервалы. Так, по численности рабочих промышленные предприятия группируются следующим образом: до 100 человек, 101—200 человек, 201—500 человек, 501—1000 человек, 1001—3000 человек, 3001—10 000 человек, 10 001 человек и выше. Неравные интервалы применяются при группировке колхозов по числу дворов, размерам валового дохода, поголовью скота и т. д. Неравные интервалы применяются также при группировке по различным признакам торговых предприятий, транспортных организаций и т. п.

Широкое применение неравных, прогрессивно увеличивающихся интервалов объясняется тем, что для большинства экономических явлений количественные изменения размера признака имеют неодинаковое значение в низших и высших по размеру признака группах. Так, если разница в 50, а тем более в 100 рабочих имеет существенное значение для мелких промышленных предприятий, на которых насчитывается всего 100—200 рабочих, то для предприятий, имеющих несколько тысяч рабочих, такая разница существенного значения не представляет. То же можно сказать о размерах валового дохода при группировке колхозов по этому признаку и т. д.

Вместе с тем иногда целесообразно применение группировок с равными интервалами. Равные интервалы применяются обычно при относительно узких пределах вариации признака и распределении единиц, близком к равномерному, например, при группировке рабочих по размерам заработной платы, маргеновских показателей и т. д.

Определение величины интервала в случае группировки с применением равных интервалов производится по формуле

$$h = \frac{x_{\max} - x_{\min}}{\text{число групп}} \quad (4.1)$$

где h — величина интервала, x_{\max} — максимальная величина признака и x_{\min} — минимальная величина признака.

Предположим, что величина месячной заработной платы рабочих промышленного предприятия колеблется в пределах от 100 до 300 руб. Необходимо произвести группировку рабочих по размеру заработной платы, образовав при этом 10 групп с равными интервалами. Величина интервала составит:

$$h = \frac{300 - 100}{10} = 20 \text{ руб.}$$

Прибавляя к минимальному значению признака (в данном случае 100 руб.) найденное значение интервала, получаем верхнюю границу первой группы: $100 + 20 = 120$ руб. Прибавляя далее величину интервала к верхней границе первой группы, получаем верхнюю границу второй группы: $120 + 20 = 140$ руб. и т. д. В ре-

зультате получим такие группы рабочих по размеру заработной платы (руб.): 100—120; 120—140; 140—160 и т. д.

Интервалы могут быть не только *замкнутыми* (с указанием как нижней, так и верхней границ), но и *открытыми* (с указанием только одной из границ — см. табл. 4.19). Открытые интервалы применяются только для крайних групп.

При количественных группировках следует обращать внимание на правильное обозначение *нижней и верхней границ* каждой группы.

Так, если к приведенным выше группам рабочих по размеру месячной заработной платы не дать специальных указаний, то возникнут трудности при составлении и чтении данной группировки. Куда, например, отнести рабочего с заработком, равным 140 руб., во вторую или в третью группу? Для устранения подобных неопределенностей требуются дополнительные указания о том, считать ли верхние границы групп «включительно» или «исключительно». Если считать верхние границы групп «включительно», то рабочие с заработком 120 руб., 140 руб и т. д. должны быть отнесены соответственно в первую, вторую и т. д. группы. Если же считать верхние группы «исключительно», то рабочие с указанными заработками должны быть отнесены во вторую, третью и т. д. группы.

Тот же результат достигается и с помощью открытых интервалов в первой и последней группах. Чтобы показать, что рабочие с заработной платой, равной верхней границе интервала, включаются в данную группу, последнюю группу следует обозначать так: «до 280 руб.». Наоборот, чтобы показать, что верхние границы интервалов не входят в данную группу, последнюю группу нужно обозначать «280 руб. и более».

Часто для придания группировкам большей определенности верхнюю границу предыдущей и нижнюю границу последующей групп обозначают различно. Если группировочный признак может принимать только целые значения, то нижняя граница последующей группы отличается от верхней границы предыдущей группы на одну целую единицу. Например, группировка колхозов по числу дворов: до 100, 101—200, 201—300, 301—500, свыше 500. Если же значения признака могут быть и дробными, то нижняя граница последующей группы может отличаться от верхней границы предыдущей группы как на целую единицу, так и на часть единицы. Например, группировка колхозов по валовому доходу на один человека (руб.): до 2; 2,01—3; 3,01—4; 4,01—6, свыше 6.

Третьей важнейшей задачей статистических группировок является исследование связей и зависимостей между признаками единиц статистической совокупности. На основе типологических и вытекающих из них структурных группировок становится возможным изучение связи между общественными явлениями и отдельными их признаками. Этот анализ связи производится с помощью группировок, в которых прослеживается, как с

изменением размеров одного из признаков меняются размеры других признаков единиц изучаемой совокупности.
 Следовательно, аналитическими называются такие группировки, которые определяют взаимосвязи между различными признаками единиц статистической совокупности.

В качестве примера приведем следующую группировку:

Таблица 4.4

ЗАВИСИМОСТЬ СРЕДНЕЙ МЕСЯЧНОЙ ЗАРАБОТНОЙ ПЛАТЫ РАБОЧИХ ПРОМЫШЛЕННОГО ПРЕДПРИЯТИЯ ОТ КВАЛИФИКАЦИИ

Группы рабочих по квалификации (разрядам)	Средняя месячная заработная плата	
	абсолютная (руб.)	в процентах к первой группе
1	115	100
2	127	110
3	138	120
4	149	130
5	175	152
6	179	156
Итого	140	122

Таблица показывает прямую зависимость средней месячной заработной платы рабочих от их квалификации: чем выше квалификация, тем выше и средняя месячная заработная плата. Заработная плата наиболее квалифицированных рабочих (6-й разряд) выше заработной платы неквалифицированных рабочих (1-й разряд) в 1,6 (округленно) раза.

Применяя к обработке аналитических группировок методы математической статистики, статистика получает возможность численной характеристики формы и силы связей взаимосвязанных признаков единиц статистической совокупности. Подробнее этот вопрос рассматривается в гл. VI и VIII.

Вторичная группировка

Статистические группировки производят главным образом на основе первичного статистического материала. Но наряду с такими группировками, которые можно назвать первичными, в статистике применяются группировки на основе уже сгруппированного материала — так называемые *вторичные группировки*. К вторичной группировке обычно прибегают в тех случаях, когда произведенные ранее группировки не удовлетворяют целям исследования либо в отношении числа групп, либо с точки зрения сопоставимости данных.

Вторичная группировка — очень важный прием статистической сводки. Вторичную группировку с большим успехом применял В. И. Ленин, например, при изучении развития капитализма в

России, при обработке итогов первой всеобщей переписи населения России 1897 г., при обработке материалов американской аграрной статистики и в ряде других исследований.

Различают два способа образования новых, обычно укрупненных групп: 1) путем изменения интервалов первичной группировки и 2) путем выделения определенной доли единиц совокупности.

Первый способ применяют для обеспечения сравнимости результатов нескольких группировок по одному и тому же признаку, но с разными границами групп. Например, колхозы двух районов разгруппированы по поголовью коров. В районе I менее 300 коров имеют 20% всех колхозов, от 300 до 499 коров — 52%, 500 коров и более — 28% колхозов. В районе II образованы такие группы (см. первичную группировку в табл. 4.5).

Таблица 4.5

ПЕРВИЧНАЯ И ВТОРИЧНАЯ ГРУППИРОВКИ КОЛХОЗОВ РАЙОНА II ПО ПОГОВЬЮ КОРОВ (на 1 января 1974 г.)

Первичная группировка		Вторичная группировка	
поголовье коров	число колхозов (в процентах к итогу)	поголовье коров	число колхозов (в процентах к итогу)
Менее 200	16	Менее 300	$16 + 24 \times 0,5 = 28$
200—399	24	300—499	$24 \times 0,5 + 42 \times 0,5 = 33$
400—599	42	500 и более	$42 \times 0,5 + 18 = 39$
600 и более	18		
Итого	100	Итого	100

Чтобы сделать данные сопоставимыми, произведем вторичную группировку колхозов района II, образован те же три группы, что и в районе I. Тогда в первую группу вторичной группировки (менее 300 голов) войдут 16% колхозов с поголовьем менее 200 голов и некоторая часть колхозов с поголовьем от 200 до 399 голов. Условно считая, что колхозы внутри этой группы распределены равномерно по всей ширине интервала (и учитывая, что верхняя граница первой группы вторичной группировки (300) делит интервал 200—399 пополам, найдем, что в первую группу нужно добавить половину колхозов, имеющих от 200 до 399 коров, т. е. 12% ($24 \times 0,5$). Остальные 12% колхозов, имеющих от 300 до 399 коров, войдут во вторую группу вторичной группировки (300—499 коров). Кроме того, в эту группу войдет половина колхозов, имеющих от 400 до 599 коров, т. е. 21% ($42 \times 0,5$). Наконец, в третью группу вторичной группировки войдет половина колхозов, имеющих от 400 до 599 коров, т. е. 21%, и все колхозы с поголовьем 600 коров и более.

Если бы по каждой группе, кроме числа колхозов, были приведены также другие сведения, например о поголовье коров по всем

колхозах данной группы, о валовом надое молока и т. д., то распределение этих показателей по новым группам следовало бы производить аналогично распределению числа колхозов.

Сопоставляя полученные по району II результаты (см. табл. 4.5) с данными по району I (20, 52 и 28), видим, что в районе I поголовье коров распределяется по колхозам более равномерно, чем в районе II, где сильнее выражена дифференциация колхозов по размеру поголовья и концентрации самого поголовья коров в крупных по этому признаку хозяйствах.

Сущность второго способа вторичной группировки, основанного на выделении определенной доли единицы совокупности, покажем на ленинском материале. Изучая экономическое положение крестьянского хозяйства пореформенной России, В. И. Ленин привлек материалы земской статистики. Но эти материалы имели ряд существенных недостатков. В частности, так как статистические учреждения проводили обследования крестьянских хозяйств по различным программам, то и группировка материалов осуществлялась на основе различных группировочных признаков. В одних губерниях крестьянские хозяйства группировались по размеру посевов, в других — по наличию рабочего скота и т. д. При этом число групп устанавливалось произвольно. Эти материалы нельзя было свести не только в целом для всей России, но и для отдельных губерний они не давали правильной группировки крестьянских хозяйств по социально-экономическим типам.

Чтобы получить правильное, с позиций экономической теории, представление о дифференциации крестьянских хозяйств, В. И. Ленин произвел вторичную группировку материалов земской статистики. При этом во всех губерниях было взято соотношение дворов, соответствующее в среднем основным классовым группам крестьянства: 50% бедняцких дворов, 30% середняцких, 20% зажиточных, т. е. в основу группировки были положены удельные веса трех социально-экономических типов крестьянского хозяйства в общем числе хозяйств. Опираясь на эту группировку, В. И. Ленин показал, что как в отдельных губерниях, так и в целом по России происходит быстрое развитие капитализма в сельском хозяйстве.

Приемы вторичной группировки, примененные В. И. Лениным, покажем на следующих данных земской статистики о распределении домохозяев Красноуфимского уезда Пермской губернии по размерам обрабатываемой земли¹.

Чтобы образовать три социально-экономических типа крестьянских хозяйств, нужно согласно условию сначала выделить из числа низших по экономической мощности групп 50% наиболее бедных хозяйств. Сюда попадают прежде всего первые две группы первичной группировки — не обрабатывающие землю и обрабатывающие до 5 дес. Они составляют 40,5% (10,2+30,3) всех хозяйств. Следовательно, до 50% недостает еще 9,5% (50,0 — 40,5), которые нужно

Группы домохозяев	% дворов	% ко всему количеству посевов	% ко всему количеству скота
не обрабатывающие земли	10,2	—	1,7
от 5 д	30,3	8,9	13,7
от 5—10 »	27,0	22,4	24,5
от 10—20 »	22,4	35,1	33,8
от 20—50 »	9,4	28,9	23,2
свыше 50 »	0,7	4,7	3,1
Всего	100	100	100

выделить из третьей группы хозяйств, обрабатывающих от 5 до 10 дес.

Переходим к расчету посевной площади для этой вновь образованной группы бедных хозяйств. Посевная площадь их равна посевной площади первых двух групп (0+8,9%) и еще какой-то части посевной площади третьей группы. Условно примем, что посевная площадь внутри каждой группы распределяется между хозяйствами равномерно, т. е. все хозяйства внутри группы имеют одинаковые посевные площади. При таком предположении можно узнать размер посевной площади из третьей группы пропорционально числу хозяйств, выделяемых из нее в низшую группу вторичной группировки. Необходимо, однако, иметь в виду, что при неравномерном распределении показателя внутри групп подобные расчеты будут несколько неточны.

Всех хозяйств в третьей группе (обрабатывающие от 5 до 10 дес.) 27,0%, а нужно выделить из нее и присоединить к первым двум, как сказано выше, только 9,5%. Эта часть составляет 0,352 (9,5 : 27,0) всех дворов третьей группы и обладает 7,9% посевной площади (22,4% × 0,352). Таким образом, вся посевная площадь несостоятельных крестьянских дворов будет равна 16,8% (0+8,9%+7,9%) посевной площади крестьянских хозяйств уезда.

Таким же образом рассчитывается и процент скота, принадлежащего низшей группе. Численность скота 50% дворов несостоятельных хозяйств складывается из поголовья скота первых двух групп (1,7%+13,7%), а также (исходя из условного предположения о равномерном распределении скота среди хозяйств каждой группы) из 0,352 поголовья скота третьей группы (24,5% × 0,352 = 8,6%). В целом это дает 24% (1,7%+13,7%+8,6%) общего количества скота всех крестьянских хозяйств уезда.

Аналогичным образом выделяется 20% зажиточных хозяйств. В эту группу войдут прежде всего дворы двух нижних в таблице групп, которые составляют 10,1% всех дворов (0,7%+9,4%). Недостающие до 20% 9,9% дворов (20,0% — 10,1%) надо взять из группы домохозяев, обрабатывающих от 10 до 20 дес.

¹ См.: Ленин В. И. Полн. собр. соч., т. 3, с. 97.

Посевная площадь для этой группы в 20% зажиточных хозяйств будет состоять из посевной площади двух последних в таблице групп (4,7% + 28,9%) и части, равной 0,442 (9,9% : 22,4%) от посевной площади группы с посевом от 10 до 20 дес., т. е. 35,1% × 0,442 = 15,5%. Всего, таким образом, посевная площадь зажиточной группы хозяйств составит 49,1% (4,7% + 28,9% + 15,5%) от посевной площади всех крестьянских хозяйств уезда.

Окончательные результаты распределения дворов, посевной площади и скота по социально-экономическим типам хозяйств представлены в следующей таблице.

Таблица 4.7
РАСПРЕДЕЛЕНИЕ ДОМОХОЗЯЕВ
КРАСНОУФИМСКОГО УЕЗДА ПЕРМСКОЙ ГУБЕРНИИ
ПО СОЦИАЛЬНО-ЭКОНОМИЧЕСКИМ ТИПАМ

	Процент дворов к итогу	Посевная площадь группы (в процентах к итогу)	Всего скота у группы (в процентах к итогу)
Низшая	50	16,8	24,0
Средняя	30	34,1	34,8
Вышая	20	49,1	41,2
Итого	100	100	100

Из таблицы видно, что пятая часть (20%) зажиточных хозяйств держит в своих руках почти половину всей посевной площади и скота, тогда как половина всех хозяйств (низшая группа) владеет лишь шестой частью посевной площади и менее чем четвертой частью скота. Таковы были вскрытые В. И. Лениным с помощью метода вторичной группировки результаты проникновения капитализма в сельское хозяйство России.

Произведенная по материалам земской статистики вторичная группировка обладает теми достоинствами, что, во-первых, вместо большого числа распыленных групп она дает всего три группы, представляющие явно выраженные социально-экономические типы, во-вторых, что только таким путем достигается сравнимость данных о разложении крестьянства в различных местностях с различными условиями.

Балансовый метод — особое место в сводке и группировке материалов статистического наблюдения занимает балансовый метод. Сущность его заключается в характеристике ресурсов изучаемого явления и их распределения (использования). Для этого составляется двойная, балансового типа, таблица. Помещаемые в таблицу данные о тех или иных явлениях излагаются так, чтобы между обеими ее частями было равновесие, баланс.

Наиболее простым видом баланса является баланс отдельного продукта, в котором соблюдается следующее балансовое равен-

ство: остаток на начало периода + поступление = расход + остаток на конец периода.

Аналогично могут быть построены балансы отдельных видов производства, баланс денежных доходов и расходов и т. п. В качестве примера приведем следующий баланс.

Таблица 4.8
БАЛАНС ТОПЛИВНО-ЭНЕРГЕТИЧЕСКИХ РЕСУРСОВ СССР ЗА 1971 г.

(выражено на условное топливо; млн. т)

Ресурсы		Распределение	
Всего	1637,5	Всего	1637,5
Производство (добыча)		Израсходовано	1309,7
Остаток на начало периода	1420,6	в том числе:	
Производства гидроэлектроэнергии	15,0	на выработку электроэнергии и теплотвергии	510,9
Поступления	44,0	на производственно-технологические и прочие нужды (включая потери при хранении и транспортировке)	798,8
Остаток на конец года	35,8	Экспорт	204,2
	122,1	Остатки на конец года	123,6

Баланс показывает, из чего складывались ресурсы и как они были использованы. На его основе можно определить, какую часть этих ресурсов составлял тот или иной их источник и какая часть ресурсов была использована для той или иной цели.

Особым видом баланса является так называемый шахматный баланс. Примером такого баланса может служить межрайонный баланс грузооборота. В нем каждый район рассматривается с двух сторон: 1) как район, вывозящий данный продукт, и 2) как район, ввозящий его.

Таблица 4.9
МЕЖРАЙОННЫЙ БАЛАНС ГРУЗОБОРОТА ПРОДУКТА К (тыс. т)

Районы вывоза	Районы ввоза				Всего вывезено	в том числе в другие районы
	А	Б	В	Г		
А	28	6	5	1	40	12
Б	7	21	2	4	34	13
В	2	8	22	6	38	16
Г	4	—	3	17	24	7
Всего вывезено в том числе из других районов	41	35	32	28	136	48
	13	14	10	11	48	

Каждая строка этого баланса показывает, в какие районы и в каком количестве был вывезен продукт из данного района, а

каждая вертикальная графа — из каких районов и сколько было ввезено продукта в данный район. Числа, находящиеся на пересечении одноименных строк и граф (т. е. 28, 21, 22 и 17), показывают размер внутрирайонного грузооборота.

Баланс позволяет изучать территориальное размещение ресурсов, соотношение между ресурсами и потребностями каждого района, связи между районами, размер встречных перевозок и т. д.

По аналогичной схеме строится межотраслевой баланс производства и потребления важнейших продуктов, в котором каждая отрасль рассматривается, с одной стороны, как производитель определенного продукта, а с другой — как потребитель продукции других отраслей. Такой баланс, дополненный системой математических уравнений, выражающих количественные соотношения между его элементами, позволяет изучать межотраслевые связи в народном хозяйстве и исследовать закономерности процесса воспроизводства.

В настоящее время советская статистика систематически составляет разнообразные балансы, в том числе балансы народного хозяйства союзных республик и в целом по СССР, балансы национального дохода, балансы труда в промышленности, сельском хозяйстве, строительстве, материальные балансы, балансы основных фондов балансы топливно-энергетических ресурсов и т. д.

Группировки как неотъемлемый элемент сводки материалов статистического наблюдения применяются и в буржуазной статистике. Однако обычно они имеют ненаучный (формальный) характер. Этот порок буржуазной статистики отмечал еще В. И. Ленин. Германскую статистику, применившую при сводке материалов группировку хозяйств на 25 групп, В. И. Ленин характеризовал как образец чиновничьей рутинности, научного хлама, бессмысленной «игры в цифирь», так как «...ни тени разумных, рациональных, научной и жизнью оправдаваемых, оснований для признания типичности такого количества таких групп не имеется»¹.

Группировка по формальным признакам часто приводит к тому, что в одну и ту же группу попадают представители разных социальных типов, разных классов, в результате чего обобщающие показатели по таким группам становятся нетипичными, искажающими действительное положение вещей, приукрашивающими картину в угоду интересам буржуазии.

Формальный, апологетический характер носят многие группировки и в современной буржуазной статистике, особенно в области социальной структуры общества. Американская официальная статистика, например, различает три социальные группы: 1) работники, получающие заработную плату или жалованье, 2) самостоятельные работники и 3) неоплачиваемые члены семей.

В первой группе наряду с собственно рабочими и служащими находятся и директора предприятий, а также высокооплачиваемые чиновники. Ко второй группе наряду с крупными капиталистами-предпринимателями относятся и мелкие фермеры, имеющие одно-двух рабочих, а также лица свободных профессий (врачи, художники и т. п.), прибегающие к услугам наемного персонала. К третьей группе относятся как члены семей предпринимателей, получающие пособия, так и члены бедных крестьянских семей и семьи ремесленников, работающие без оплаты. Таким образом, все группы оказываются неоднородными в социальном отношении. Естественно, что такая группировка создает неправильное представление о социальной структуре общества. Выводимые же на основе этой группировки обобщающие показатели, например средний уровень заработной платы, создают фальшивую картину пролетариата и равенства.

¹ Ленин В. И. Полн. собр. соч., т. 27, с. 208.

Типичным образом формалистического решения вопроса о классовых группировках является и известная профессиональная классификация населения А. Мазара, принятая Бюро ценов министерства торговли США и используемая для обобщения результатов переписей населения. По этой классификации все люди занятият делятся на 4 группы, внутри которых выделены 11 «профессиональных» подгруппы: 1. «Белые воротнички»: 1) специалисты; 2) управляющие, чиновники и собственники; 3) контролеры и т. п. работники; 4) торговые работники. II. «Синие воротнички»: 1) мастера, старшие и квалифицированные рабочие; 2) подквалифицированные рабочие; 3) неквалифицированные рабочие (в том числе занятые на фермах); III. Работники сферы услуг: 1) домашняя прислуга; 2) работники услуг; IV. Фермеры: 1) фермеры и управляющие фермами; 2) рабочие, занятые на фермах.

В настоящее время, по сравнению с началом XX века, доля группы «белых воротничков» в численности занятого населения США увеличилась, а доля «синих воротничков» почти не изменилась. Отсюда буржуазные социологи делают вывод о коренной эволюции классов, о преобладании в стране «среднего класса», о вытеснении физического труда.

Основой для таких выводов нет достаточных оснований. Во-первых, непереносимое различие «белых воротничков» с работниками умственного труда, а «синих воротничков» — с рабочим классом. Значительную часть «белых воротничков» составляют работники торговли, многие из которых (продавцы, упаковщики и т. д.) занимаются физическим трудом. С другой стороны, к «работникам сферы воротничков» относят не только наемных рабочих, но и лиц, имеющих «собственное дело», т. е. мелких собственников, фактически мелкую городскую буржуазию.

Наконец, если группа контрольных работников и группа неквалифицированных и старых рабочих, занятых на фермах, могут быть признаны в совокупности социально однородными, то остальные девять групп социально совершенно неоднородны. Первая группа («специалисты») включает как лиц наемного труда, так и лиц, имеющих собственное «дело» (юридические фирмы, лечебницы и т. п.), а представители не только средней, но и крупной буржуазии; вторая группа («управляющие, чиновники и собственники») также представляет чрезвычайно неоднородную категорию — от служащих среднего звена государственных учреждений и частных компаний до владельцев крупных предприятий и банков; то же самое можно сказать и об остальных группах.

Наконец «профессиональные» группы, возрастание удельного веса одних и уменьшение других может представлять интерес для характеристики занятий населения, но не характеризует классовый состав населения и классовые сдвиги. Прямая принадлежность к классовой структуре общества на основе «профессиональных» групп, без учета отношения людей к средствам производства имеет целью скрыть действительные классовые взаимоотношения и прежде всего заглушить

классовые сдвиги и обострять классовые противоречия.

Выводы о классовых сдвиги на основе «профессиональных» групп, без учета отношения людей к средствам производства имеет целью скрыть действительные классовые взаимоотношения и прежде всего заглушить

противоположность, антагонистичность между эксплуататорами и эксплуатируемым.

В учебной литературе буржуазные статистики либо вовсе не рассматривают вопросы группировок, либо излагают их с чисто формальной точки зрения. Так, в учебнике английского статистика А. Боули «Элементы статистики» вопросы группировки статистического материала сведены к технике получения итогов по ряду чисто формальных признаков.

3. СТАТИСТИЧЕСКИЕ ТАБЛИЦЫ

Общее понятие о статистических таблицах

После того как в результате сводки материала статистического наблюдения сгруппирован, он, как правило, представляется в виде таблиц. Табличная форма изложения

результатов группировки часто бывает ярче и красноречивее многословных рассуждений.

Например, текстовое изложение данных об уровне образования населения СССР в 1959 и 1973 гг. выглядело бы следующим образом. По данным переписи 1959 г., на 1000 человек в возрасте 10 лет и старше приходилось населения, имеющего высшее и среднее (полное и неполное) образование, — 361 человек, в том числе имеющих высшее законченное образование — 23, высшее незаконченное — 11, среднее специальное — 48, среднее общее — 61 и неполное среднее — 218 человек.

В начале 1973 г. на 1000 человек в возрасте 10 лет и старше приходилось населения, имеющего высшее и среднее (полное и неполное) образование, — 522 человека, в том числе: имеющих высшее законченное образование — 49, высшее незаконченное — 14, среднее специальное — 78, среднее общее — 135, неполное среднее — 246 человек.

Этот сложный и трудный для восприятия текст может быть заменен следующей таблицей.

УРОВЕНЬ ОБРАЗОВАНИЯ НАСЕЛЕНИЯ СССР
(на начало года)

	1959 г.	1973 г.
На 1000 человек населения в возрасте 10 лет и старше приходится лиц, имеющих образование:		
высшее и среднее (полное и неполное)	361	522
в том числе:		
высшее законченное	23	49
высшее незаконченное	11	14
среднее специальное	48	78
среднее общее	61	135
неполное среднее	218	246

Таким образом, статистическая таблица представляет собой форму наиболее рационального, наглядного и систематизированного изложения числовых результатов сводки и обработки статистических материалов.

Достоинствами статистических таблиц являются выразительность, наглядность и компактность содержащихся в них данных. В таблицах отсутствуют повторения одних и тех же сведений об аналогичных показателях, имеющиеся в текстовом изложении. Вместе с тем табличное изложение цифровых характеристик общественных явлений и процессов обладает исключительной наглядностью. Если, например, при текстовом изложении сопоставление численности населения СССР, имеющего высшее и среднее образование на начало 1959 г. и на начало 1973 г., требует значительного умственного и зрительного напряжения, то в таблице эти показатели легко обозримы и сопоставимы.

Благодаря своим достоинствам статистические таблицы прочно вошли в практику экономических исследований. История свидетельствует, что практика применения статистических таблиц имеет большую давность. Еще в XVIII в. английский ученый И. К. Кирилов представил экономико-статистические сведения в табличной форме. Заслугой земских статистиков является широкое применение и развитие табличного метода. Земский статистик А. П. Шанин является создателем комбинационных таблиц.

Большое значение статистическим таблицам придавал В. И. Ленин, широко применявший их в своих работах.

Внешне статистическая таблица представляет определенную комбинацию вертикальных граф и горизонтальных строк, в которых располагаются числа. Значение показателей в каждой строке и в каждой графе поясняются заголовками, помещаемыми в левой части (наименования строк) и в верхней части таблицы (наименования граф). Кроме того, каждая статистическая таблица должна иметь общий заголовок.

Незаполненная цифрами таблица, имеющая лишь общий, боковой и верхние заголовки, носит название макета таблицы.

Назначение статистической таблицы — характеристика общественных явлений и процессов, поэтому ее можно рассматривать как форму логического предложения, имеющего статистическое содержание и статистическое сказуемое.

Подлежащим статистической таблицы является та статистическая совокупность, те объекты (или их части, группы), которые характеризуются числовыми показателями. Сказуемым статистической таблицы являются все те показатели, которые характеризуют значимую совокупность.

Возьмем, например, табл. 4.10. Подлежащим в ней является население, имеющее высшее и среднее образование, с распределением по видам образования, а сказуемым — число лиц, имеющих тот или иной уровень образования в расчете на 1000 человек населения в возрасте 10 лет и старше в 1959 и 1973 гг.

Обычно в таблицах боковые заголовки используются для обозначения подлежащего, а верхние — для обозначения статистического сказуемого. Но нередко обозначения подлежащего и сказуемого меняются местами.

То или иное расположение подлежащего и сказуемого в таблице является техническим вопросом и подчинено задаче достижения большей наглядности и компактности.

Виды статистических таблиц

В процессе экономических исследований применяются различные виды статистических таблиц. Они отличаются различным числом характеризующих в них совокупностей, различным строением и соотношением подлежащего и сказуемого и т. д. В советской экономической литературе применяются разные принципы классификации таблиц. Наиболее распространенной является классификация, согласно которой вид статистической таблицы определяется системой разработки подлежащего.

В зависимости от строения подлежащего различают три вида статистических таблиц: простые, групповые и комбинационные.

Простыми называются такие статистические таблицы, в подлежащем которых содержится простой перечень каких-либо объектов либо территориальных единиц, подвергнутых статистическому наблюдению. В соответствии с этим простые таблицы могут быть перечневыми и территориальными. Примером простой перечневой таблицы может служить таблица 4.11, в подлежащем которой содержится перечень некоторых видов промышленной продукции.

Производство некоторых видов промышленной продукции в СССР

Таблица 4.11

	1970 г.	1973 г.	1973 г. в процентах к 1970 г.
Автомобили легковые — тыс. шт.	344	917	267
Холодильники бытовые — тыс. шт.	4 140	5 423	131
Ткани всех видов — млн. м ²	8 852	9 676	109
Верхний трикотаж — млн. шт.	415	459	111
Обувь кожаная — млн. пар	676	667	99

Примером простой территориальной таблицы является табл. 4.12. Таблица остается простой и в том случае, если в ней одновременно дается и перечень единиц объекта и наименование территории.

Групповыми называются статистические таблицы, подлежащее которых содержит группировку единиц изучаемого объекта по одному существенно признаку. Простейшим видом групповых таблиц являются таблицы, в которых представлены ряды распределения (см. ниже). Более сложным видом групповых таблиц являются такие, в которых сказуемое содержит наряду с числом единиц, входящих в ту или иную группу, ряд других показателей, относящихся к характеристике групп подлежащего (см. табл. 4.7; 4.14).

Таблица 4.12

Территория и население земного шара на середину 1973 г.

	Территория (млн. км ²)	Население (млн. человек)	Плотность населения (человек на 1 км ²)
Весь мир *	135,8*	3 850	28
Европа	10,5	660	63
Азия	44,4	2 250	51
Африка	30,3	375	12
Америка	42,1	544	13
Австралия и Океания	8,5	21	2

* Без Антарктиды (14,1 млн. км²).

Комбинационными называются статистические таблицы, в подлежащем которых группы единиц, образованные по одному признаку, делятся затем на подгруппы по одному или нескольким другим признакам. Простейшим видом комбинационных таблиц являются таблицы, показывающие только распределение единиц совокупности по двум или более признакам. Например, таблицы, показывающие распределение населения одновременно по полу и возрасту либо по занятиям и уровню образования и т. д. (см. табл. 4.2).

Ниже приводится комбинационная таблица, показывающая ввод в действие жилых домов с их распределением на две группы по типу населенных пунктов: 1) города и поселки городского типа, 2) населенные пункты в сельских местностях. Каждая группа делится далее на подгруппы в зависимости от того, за чей счет были построены жилые дома, т. е. по второму признаку — источнику финансирования строительства.

Таблица 4.13

Ввод в действие жилых домов в СССР в 1973 г.

(млн. м² общей жилой площади)

	Всего	в том числе	
		в городах и поселках городского типа	в сельских местностях
государственными и кооперативными предприятиями и организациями, колхозами и населением	110,3	77,6	32,7
а том числе:			
государственными и кооперативными предприятиями и организациями и жилищно-кооперативной экономикой и служащими за свой счет и с помощью государственного кредита	82,9	70,7	12,2
в колхозах (колхозами, колхозниками и местной интеллигенцией)	13,2	6,9	6,3
	14,2	—	14,2

Более сложными являются таблицы, содержащие комбинационную группировку в подлежащем и ряд показателей в сказуемом. Приведем пример макета такой таблицы.

Таблица 4.14 (макет)

ЗАВИСИМОСТЬ ТРУДОЕМКОСТИ И СЕБЕСТОИМОСТИ
1 т МОЛОКА ОТ РАЗМЕРА ПОГЛОВОЙ КОРОВЫ
И УДОЯ МОЛОКА ОТ 1 КОРОВЫ

Группы коров по поголовью (голов)	Подгруппы по удою молока от 1 коровы (кг)	Число колхозов	Прямые затраты труда на 1 т молока (чел.- дней)	Себестоимость 1 т молока (руб.)
А	Б	1	2	3
I. Менее 400	менее 2000 2000 и более			
	Итого I группа			
II. 401 и более	менее 2000 2000 и более			
	Итого II группа			
Все колхозы	менее 2000 2000 и более			
	Всего			

Групповые и комбинационные таблицы дают глубокую и всестороннюю характеристику общественных явлений и процессов, обогащая экономический анализ.

В. И. Ленин отводил групповым и комбинационным таблицам большую роль в экономических исследованиях. Анализируя материалы подворной переписи Красносельского уезда, проведенной земскими статистиками, В. И. Ленин писал: «Все дело только за обработкой этих превосходных данных... обработка подворных данных должна дать как можно больше, как можно рациональнее и детальнее составленных групповых и комбинационных таблиц для отдельного изучения *всех* наметившихся или намечающихся (это не менее важно) — в жизни типов хозяйств. Без разносторонних и рационально составленных групповых и комбинационных таблиц богатейшие подворные данные прямо-таки пропадут»¹.

В. И. Ленин подчеркивал важное значение комбинационных таблиц в исследовании капиталистического земледелия, отмечая, что научное значение разнообразных комбинационных таблиц, учитывающих гигантское разнообразие форм подчинения земледелия и земледельца рынку, было бы громадно. Эти таблицы, по мнению В. И. Ленина, внесли бы целый переворот в науку об экономике земледелия².

¹ Ленин В. И. Полн. собр. соч., т. 24, с. 276—277.

² См.: там же, т. 24, с. 281.

Разработка сказуемого таблицы

Показатели сказуемого статистической таблицы могут быть представлены либо в *свернутом виде*, либо в *развернутом виде*, когда они подразделяются на отдельные составные части, группы. Свернутое сказуемое содержит, например, табл. 4.3, 4.4, 4.6, 4.7 и др. Развернутое сказуемое содержит приведенные ниже макеты таблиц, в которых один из показателей в сказуемом дается с распределением на группы. Развернутое сказуемое может иметь простую и сложную разработку. При простой разработке составные части показателя в сказуемом не имеют деления на подгруппы. Например:

Таблица 4.15

ТЕРРИТОРИЯ И НАСЕЛЕНИЕ СОЮЗНЫХ РЕСПУБЛИК
(на 15 января 1970 г.)

Союзные республики	Территория (тыс. км²)	Численность населения (тыс. человек — всего)	в том числе		Из общей численности населения *		Плотность населения (человек на 1 км²)
			городское	сельское	мужчин	женщин	
А	1	2	3	4	5	6	7

При сложной разработке составные части показателя в сказуемом делятся на подгруппы:

Таблица 4.16

ТЕРРИТОРИЯ И НАСЕЛЕНИЕ СОЮЗНЫХ РЕСПУБЛИК
(на 15 января 1970 г.)

Союзные республики	Территория (тыс. км²)	Численность населения (тыс. человек)		в том числе							
				городское				сельское			
				мужчины	женщины	всего	мужчины	женщины	всего	мужчины	женщины
А	1	2	3	4	5	6	7	8	9	10	11

Сложная разработка показателей сказуемого дает более полную характеристику объекта, но приводит к увеличению размеров таблиц, что уменьшает наглядность и затрудняет пользование ими. Вот почему при составлении плана сводки нужно учитывать как положительные, так и отрицательные моменты сложной разработки показателей сказуемого.

Эти отрицательные моменты в равной мере относятся и к подложному. Именно поэтому переносят отдельные части подлежащего из боковых заголовков в верхние для упрощения таблицы.

Показатели сказуемого могут характеризовать состояние объекта на какой-либо один момент времени (см. табл. 4.12) или итоги протекшего процесса за один отрезок времени (см. табл. 4.13).

Такие таблицы называются *статическими*. Если же показатели сменяемого характеризуют изменение объекта во времени, то такие таблицы называются *динамическими* (см. табл. 4.10 и 4.11).

Оформление статистических таблиц

Чтобы таблица достигла наибольшей выразительности, необходимо при ее оформлении руководствоваться следующими правилами.

Форма статистической таблицы должна быть согласована с ранее существующими таблицами для обеспечения возможности сравнения данных за ряд отрезков времени.

Заглавие таблицы должно кратко и точно характеризовать основное ее содержание. Требование точности, четкости и ясности относится и к заголовкам строк и граф. Если заголовок недостаточно подробно сформулирован, то можно сделать примечания к нему (см. табл. 4.12).

В таблице (в общем заголовке или в заголовках строк и граф) должно быть указано, в какой территории и к какому периоду или моменту времени относятся приводимые в ней данные, а также характер приводимых данных (планоые, фактические, расчетные) и единицы измерения.

Для удобства чтения таблица должна быть по возможности небольшой. Взамен одной громоздкой таблицы целесообразно строить две-три меньших размеров.

Если число граф (или строк) в таблице велико, то они должны нумероваться. При этом графы, в которых содержится перечень объектов или их группы, обозначаются заглавными буквами алфавита, а графы, содержащие показатели сменяемого, — арабскими цифрами (см. табл. 4.14 и 4.15 на с. 102—103).

В случае отсутствия сведений в соответствующей клетке таблицы записывается «Нет сведений» или проставляются точки (...). Отсутствие факта обозначается прочерком (—). Число 0,0 ставится в тех случаях, когда величина показателя не превышает 0,05 (см. табл. 4.1).

Таблицы, как правило, должны быть замкнутыми, т. е. иметь итоговые результаты (в целом, по группам и подгруппам).

4. ГРАФИЧЕСКИЕ СПОСОБЫ ИЗОБРАЖЕНИЯ СТАТИСТИЧЕСКИХ ДАННЫХ

Сущность и значение графиков

Статистические данные, т. е. результаты сводки и дальнейшей обработки собранных материалов, наряду с табличной формой могут быть изображены также в виде графиков.

Графиками в статистике называются основные изображения числовых величин и их соотношений в виде различных геометрических образов: точек, линий, фигур и т. п.

Главное достоинство графиков — наглядность. При графическом изображении статистические данные привлекают внимание, производят более яркое и живое впечатление, становятся более доход-

ными и запоминающимися. Поэтому графики используются в агитационной и пропагандистской работе для популяризации статистических данных. Так, еще в 1919 г. в подписанном В. И. Лениным постановлении правительства предусматривалось создание на площадях, в театрах и в других местах сосредоточения населения сети витрин с изображением статистических данных в виде диаграмм, картограмм и т. д.

В настоящее время графики широко используются для наглядной иллюстрации статистических данных, содержащихся в Программе КПСС, в материалах съездов партии и Пленумов ЦК КПСС. В газетах и журналах, на выставках и уличных стендах, на предпринятиях — всюду можно найти графики, наглядно изображающие наши успехи и перспективы развития экономики и культуры.

Исключительно велика роль графиков как средств обобщения и анализа статистических данных.

Графики дают целостную картину явлений и процессов, обобщающее представление о них и помогают осмыслить статистический материал. При графическом изображении статистических данных становятся особенно отчетливыми и наглядными взаимная связь между явлениями и процессами общественной жизни, основные тенденции их развития, степень распространения их в пространстве.

Важное значение графиков подчеркнул В. И. Ленин. В книге «Шаг вперед, два шага назад», анализируя работу II съезда РСДРП, В. И. Ленин отмечал, что разделение делегатов по различным обсуждавшимся на съезде вопросам дает единственную в своем роде картину внутренней борьбы в партии. «Чтобы сделать эту картину наглядной», — писал В. И. Ленин, — «чтобы получить настоящую картину, а не грудку бессвязных, дробных, изолированных фактов и фактиков... я решил попытаться изобразить все основные типы «разделений» нашего съезда в виде диаграмм. Такой прием покажется, наверное, странным очень и очень многим, но в то же время, можно ли найти другой способ изложения, действительно обобщающего и подводящего итоги, возможно более полного и наиболее точного?»¹

Графики прочно вошли в практику экономической работы. Умея читать и строить графики необходимо любому экономисту.

Каждый график состоит из графического образа и вспомогательных элементов. Графический образ — это совокупность точек, линий и фигур, с помощью которых изображены статистические данные. Так, на рис. 4.1 графический образ представляет собой ряд столбиков, на рис. 4.6 — ряд квадратов и т. д.

Если бы график состоял только из точек, линий или фигур, то трудно было бы понять, что они изображают. Поэтому графический образ обязательно должен быть разъяснен с помощью пояснитель-

¹ Ленин В. И. Полн. собр. соч., т. 8, с. 321—322.

ных подписей и условных обозначений, являющихся вспомогательными элементами графика. К ним относятся: 1) общий заголовок графика; 2) словесные пояснения условных знаков и смысла отдельных элементов графического образа; 3) оси координат, шкалы с масштабами и числовые сетки; 4) числовые данные, которые дополняют или уточняют величину нанесенных на график показателей.

Вспомогательные элементы делают возможным чтение графика, его понимание и использование. Так, на рис. 4.1 только из общего заголовка видно, что речь идет о средней продолжительности жизни населения нашей страны. Пояснительные надписи, помещенные под столбиками, указывают, к какой части населения и к какому периоду относятся данные о средней продолжительности жизни.

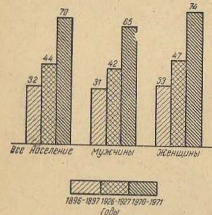


Рис. 4.1. Средняя продолжительность жизни населения СССР и Российской Республики (в годах)

Последний способ применяется обычно в тех случаях, когда на графике недостаточно места, а пояснения длинные.

Оси координат, шкалы с масштабами и числовые сетки необходимы для точного построения и чтения графика.

Шкалой называется линия, отдельные точки которой могут быть прочитаны как определенные числа. Шкалы могут быть прямыми или криволинейными (круговыми), равномерными (арифметическими) и неравномерными. Если равным графическим интервалам (отрезкам шкалы) соответствуют равные числовые интервалы, то шкала является равномерной (рис. 4.2). Если же равным графическим интервалам соответствуют неравные числовые интервалы, то шкала будет неравномерной. Примером такой шкалы может служить логарифмическая шкала (см. рис. 4.2). На ней началом отсчета является единица ($\lg 1 = 0$), а расстояния от начала отсчета до точек пропорциональны не числам, а их логарифмам. Равным графическим интервалам здесь соответствуют равные интервалы (разности) логарифмов чисел, но неравные интервалы самих чисел.

Масштабом равномерной шкалы называется длина отрезка, принятого за числовую единицу того или иного наименования. Так, на рис. 4.1 за 1 год принят отрезок 0,5 мм. Этот масштаб можно выразить по-другому: 1 мм соответствует 2 года жизни. Чем длиннее отрезок, принятый за числовую единицу, тем крупнее масштаб (на рис. 4.2 первый масштаб в 3 раза крупнее, чем второй).

На одной линии или на двух параллельных линиях можно построить две связанные между собой шкалы. Такие шкалы называются сопряженными. Чаще всего одна из сопряженных шкал используется для изображения и отсчета абсолютных величин, а другая — для отсчета процентов (см. рис. 5.1).

Если на координатном поле через деления горизонтальной и вертикальной шкал провести прямые линии, то образуется числовая сетка, которая облегчает построение и чтение графика (см. рис. 5.1).

При наличии густой сетки числовые значения изображенных показателей на графике можно не наносить, однако они должны быть приведены в тексте.

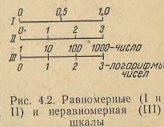
Шкала, по которой отсчитываются уровни явлений, должна, как правило, начинаться с нуля (0). Числа на шкале представляются равномерно, через определенные интервалы. Последнее число должно несколько превышать максимальный уровень, отсчет которого будет производиться по этой шкале. Над числами шкалы обычно указываются единицы их измерения (см. рис. 4.3 и 4.4).

Числовая сетка, как правило, должна иметь базовую линию. Обычно базовой линией является горизонтальная линия (ось абсцисс в системе прямоугольных координат). Если на числовой сетке строится линейный график (ломаная кривая), занимающий лишь верхнюю часть сетки, то нижняя ее часть может быть исключена путем разрыва шкалы и сетки (см. рис. 8.1 и 8.4).

Исключение части сетки дает возможность при том же размере графика укрупнять масштаб.

Статистические графики различаются по содержанию и по способу построения. По содержанию изображаемых статистических показателей графики делятся на следующие виды: 1) графики сравнения одноименных показателей, относящихся к разным объектам или территориям; 2) графики вариационных рядов; 3) графики структуры (состава); 4) графики взаимозависимых показателей; 5) графики динамики (изменения во времени); 6) графики выполнения плана.

В этой главе рассматривается построение и использование графиков сравнения и вариационных рядов; остальные виды графиков (графики структуры, взаимозависимых показателей, динамики и выполнения плана) рассматриваются в других главах, в которых говорится о соответствующих статистических показателях.



В ряде случаев графики по своему содержанию являются комбинированными, т. е. одновременно решают несколько задач, например изображают структуру и ее изменение во времени (структурные сдвиги). Комбинированным является и график, приведенный на рис. 4.1, на котором одновременно с изображением роста средней продолжительности жизни дается сравнение ее величины у мужчин и женщин.

По способу построения графики делятся на *диаграммы*, *картодиаграммы* и *картограммы*. *Диаграмма* представляет собой чертёж, на котором статистические данные представлены при помощи *геометрических фигур или знаков*, а территория, к которой относятся эти данные, указана только словесно (рис. 4.1, 4.3). Если *диаграмма* наложена на карту или план территории, к которой относятся изображенные показатели, то график называется *картодиаграммой* (рис. 5.7). Если же величина показателей изображена путем штриховки или раскраски соответствующей территории на карте или плане, то график называется *картограммой* (рис. 5.8).

Каждый из этих видов графиков имеет ряд разновидностей. Так, диаграммы в зависимости от формы графического образа могут быть точечными, линейными, плоскостными, пространственными и фигурными. Применяются и комбинированные диаграммы.

Графики сравнения

Для сравнения одноименных показателей, характеризующих разные объекты или территории, могут быть использованы различные виды диаграмм. Наиболее наглядными являются столбиковые диаграммы. Каждое значение сравниваемого показателя изображается в этом случае в виде вертикального прямоугольного столбика. Основания всех столбиков располагаются на горизонтальной базовой линии. Ширина столбиков берется произвольная, но обязательно одинаковая для всех сравниваемых величин. Высота каждого столбика пропорциональна величине изображаемого показателя, что достигается принятием одинакового для всех столбиков масштаба.

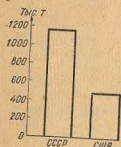


Рис. 4.3. Производство животного масла в СССР и США в 1972 г.

равна 47 мм (4 мм \times 11,76), высота второго столбика (США) должна быть $\times 5,03$; см. рис. 4.3).

В столбиковых диаграммах основой сравнения показателей является высота столбиков над горизонтальной базовой линией. Поэтому разрыв шкалы (сетки) не допускается.

Если базовая линия расположена вертикально, а столбики — горизонтально, то диаграмма называется *ленточной* (*полосовой*).

Если при этом каждый объект или территория могут характеризоваться двумя показателями (особенно противоположными по смыслу), целесообразно построить двустороннюю полосовую диаграмму (см. рис. 4.4).

Диаграммы, предназначенные для популяризации, иногда строятся в виде стандартных фигур — рисунков, характерных для изображаемых показателей, что делает диаграмму более выразительной и привлекает к ней внимание. Такие диаграммы называются *фигурными*, или *изобразительными*. Каждая фигура имеет одинаковый размер и принимается за определенную величину изображаемого показателя. Так, при изображении численности населения одна фигура, изображающая силуэт человека, может быть при-

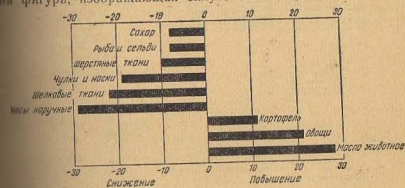


Рис. 4.4. Изменение государственных розничных цен на отдельные товары в 1972 г. по сравнению с 1960 г. (процентов)

нито за 100 тыс. человек или за 1 млн. человек, при изображении жилищного строительства одно здание может соответствовать 1 млн. м² общей площади жилищ и т. п. (см. рис. 4.5).

Недостаток фигурных диаграмм заключается в том, что во многих случаях приходится либо округлять изображаемые показатели, либо изображать, кроме целых фигур, их части, размер которых не сразу оценить трудно.

Иногда разница между наименьшими и наибольшими значениями сравниваемого показателя настолько велика, что установление подходящего масштаба для столбиков или полос оказывается затруднительным. Пусть, например, нужно изобразить графически следующие данные о численности населения частей света на середину 1973 г. (млн. человек): Азия — 2250, Европа — 660, Америка — 544, Африка — 375, Австралия и Океания — 21.

Если попытаться изобразить эти данные в виде столбиковой диаграммы, то высота наибольшего столбика, изображающего население Азии, должна быть в 107 раз больше, чем высота наименьшего столбика, изображающего население Австралии и Океании (1973 г. 21).

В случаях значительного размаха колебаний показателя вместо столбиковой целесообразно применить плоскостную (двухмерную)

диаграмму — квадратную или круговую. Принцип построения этих диаграмм заключается в том, что величины сравниваемых показателей изображаются площадями квадратов или кругов. Иными словами, площади квадратов (кругов) должны быть пропорциональны величинам изображаемых показателей. Но площади квадратов

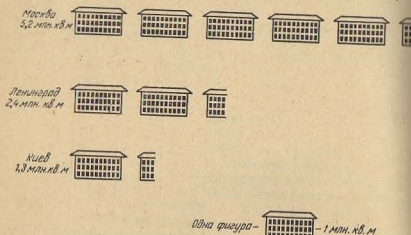


Рис. 4.5. Жилищное строительство в Москве, Ленинграде и Киеве в 1972 г. (общая (полезная) площадь построенных квартир; млн. кв. м)

(кругов) пропорциональны квадратам их сторон (радиусов). Следовательно, стороны квадратов или радиусы кругов должны быть пропорциональны корням квадратным из величин изображаемых показателей.

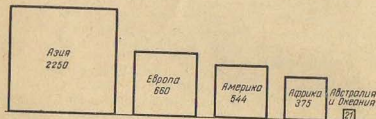


Рис. 4.6. Численность населения частей света на середину 1973 г. (млн. человек)

Приняв в нашем примере численность населения Австралии и Океании за 1, найдем, что площади квадратов должны относиться как $107:31:26:18:1$. Извлекая квадратные корни из этих чисел, получим соотношение сторон квадратов (округленно): $10:6:5:4:1$.

Расположив квадраты на горизонтальной базовой линии, получим диаграмму, изображенную на рис. 4.6.

Масштаб в квадратных и круговых диаграммах выражается площадью, принятой за единицу того или иного наименования. Так, если в нашем примере сторону квадрата, изображающего население Азии, принять равной 50 мм, то 2250 млн. человек будут соответствовать на диаграмме 2500 мм^2 , т. е. 1 млн. человек — примерно $1,1 \text{ мм}^2$, или 900 тыс. человек в 1 мм^2 . Воспринимаются такие масштабы труднее, чем линейные, поэтому их можно не применять. Вместо этого на каждой фигуре следует указать величину изображаемого показателя.

Соотношение площадей квадратов или кругов на глаз оценить значительно труднее, чем, например, высоты столбиков. Поэтому двучисленные (двухмерные) диаграммы сравнения употребляются реже, чем одномерные. Еще труднее оценить соотношение объемов, а тем более с чем объемные (пространственные) диаграммы используются еще реже.

СТАТИСТИЧЕСКИЕ РЯДЫ

Понятие о статистических рядах и их видах

В результате обработки и систематизации первичных статистических материалов получают ряды цифровых показателей, которые характеризуют отдельные стороны изучаемых явлений или процессов либо их изменение во времени. Эти ряды называются статистическими.

По своему содержанию статистические ряды делятся на два вида: ряды распределения и ряды динамики.

Ряды распределения называются рядами, характеризующие распределение единиц совокупности по какому-либо одному признаку, разновидности которого расположены в определенном порядке. Ряды динамики называются рядами, которые характеризуют изменение размеров общественных явлений во времени (эти ряды рассматриваются в гл. IX).

Построение рядов распределения вытекает из принципов статистической группировки. Если при группировке данных ограничиться только одним показателем о количестве единиц в каждой группе, то получим статистический ряд, содержащий два элемента: наименования групп и их численность. Такой статистический ряд и является рядом распределения.

Ряды распределения могут быть образованы либо по атрибутивным, либо по количественным признакам. Соответственно этому различают два их вида: атрибутивные и вариационные ряды.

При группировке по атрибутивным признакам, т. е. признакам, которые характеризуют свойство, качество данного явления и не имеют количественного выражения, получают данные о том, сколько единиц обладает той или иной разновидностью атрибутивного признака. Например, сколько среди работающих на предприятии рабочих, служащих, инженерно-технических работников, сколько среди рабочих токарей, фрезеровщиков и т. д.

Построение атрибутивных рядов распределения не представляет сложности: образуется столько групп, сколько вариантов атрибутивного признака имеет изучаемое явление. Примером атрибутивного ряда распределения может служить распределение населения СССР по классам (см. табл. 4.1), на городское и сельское, распределение студентов по видам обучения (дневное, вечернее, заочное) и т. д.

Обработка атрибутивных рядов распределения включает богатый арсенал показателей (относительные и средние величины, показатели вариации, показатели силы связи и др.), которые излагаются ниже.

Построение вариационных рядов отличается от построения атрибутивных рядов распределения тем, что в этом случае принимается во внимание количественная характеристика варьирующего признака. Предположим, что имеются такие данные о числе живущих в квартирах одного из жилых домов: 3, 2, 5, 4, 6, 5, 3, 2, 4, 3, 4, 1, 2, 5, 5, 2, 3, 3, 1, 4, 4, 2, 2, 3, 2, 3, 4, 4, 5, 7, 6. Вместо того чтобы иметь дело со сложной записью, преобразуем ее, подсчитав, в скольких квартирах живут по 1, 2, 3 и т. д. человека.

РАСПРЕДЕЛЕНИЕ КВАРТИР ЖИЛОГО ДОМА
ПО ЧИСЛУ ЖИВУЩИХ В НИХ ЛИЦ

Число квартир	Число лиц, живущих в квартирах							Итого
	1	2	3	4	5	6	7	
Число квартир	2	6	7	7	5	2	1	30

В новой записи имеется вариация числа живущих в квартире, т. е. вариация численного значения изучаемого признака, и, кроме того, видно, сколько раз повторяется каждый из этих вариантов, как часто имеет место явление. Это и есть вариационный ряд.

Таким образом, вариационный ряд представляет упорядоченный ряд изменяющихся размеров количественного признака и численности единиц, имеющих данное значение признака. Числовые значения размеров количественного признака называются вариантами, а соответствующие этим вариантам численности — частотами.

Варианты могут выражаться числами положительными и отрицательными, абсолютными и относительными. Примером первого вида может служить распределение рабочих по возрасту, по заработной плате, предприятий — по числу рабочих, по объему продукции и т. д. Если же группировать предприятия по финансовым результатам хозяйственной деятельности, то варианты могут быть как положительными (прибыль), так и отрицательными (убыток) числами. Примером вариантов, выраженных относительными числами, может служить распределение предприятий по проценту выполнения плана производства продукции, производительности труда и т. д.

Частоты могут быть выражены как в абсолютных величинах, т. е. числом каких-либо единиц, так и в относительных величинах в виде долей или процентов к итогу. Частоты, выраженные в виде относительных величин, называются частотами. Частоты всегда являются положительными числами, так как, показывая, сколько раз встречается тот или иной вариант, они по своей природе не могут быть менее нуля.

Сумма частот вариационного ряда называется его объемом. Вариация признака может быть дискретной и непрерывной. Если отдельные значения вариантов могут отличаться друг от друга только на некоторую определенную величину (обычно единицу), то вариация является дискретной. Иными словами, варианты дискретного признака обычно выражаются целыми числами. Примерами дискретной вариации могут служить тарифные размеры рабочих, число обслуживаемых одним рабочим станков, число бригад в колхозе, число живущих в квартире и т. д.

Рядом с дискретной вариацией признака является, например, следующий вариационный ряд.

Таблица 4.18

УСТРОЙСТВО СЕМЕЙ ПО ЧИСЛУ СОВМЕСТНО ПРОЖИВАЮЩИХ
ЧЛЕНОВ СЕМЬИ В СССР В 1970 г.

Число совместно проживающих членов семьи (человек)	Число семей (в процентах к итогу)	Число совместно проживающих членов семьи (человек)	Число семей (в процентах к итогу)
2	25,4	7	2,8
3	26,2	8	1,5
4	24,1	9	0,8
5	12,6	10 и более	0,7
6	5,9	Итого	100

Помимо дискретной бывают вариации другого рода, когда варианты в определенных пределах могут иметь любое (целое и дробное) значение. Например, возраст людей может измеряться не только в годах, но и в месяцах, днях и даже часах, заработная плата — в рублях, копейках и даже долях копеек и т. д. Такая вариация называется непрерывной, а вариационные ряды, образованные на основе таких вариантов, называются рядами с непрерывной вариацией признака.

В случае непрерывного варьирования варианты обычно группируются в интервалы, а частоты относятся не к отдельному значению признака, как в дискретных вариационных рядах, а ко всему интервалу (см., например, табл. 4.2). Поэтому непрерывные вариационные ряды часто называют интервальными. Если интервал во всем ряде сохраняют одну и ту же величину, то они называются равными, а вариационный ряд — рядом с равными интервалами.

Часто и ряды с дискретной вариацией признака строятся путем образования групп.

Таблица 4.19

ГРУППИРОВКА КОЛХОЗОВ ПО ПОГОЛОВЬЮ КОРОВ на 1 ЯНВАРЯ 1973 г.

Колхозы, имеющие коров	Число колхозов (в процентах ко всем колхозам)
Менее 100	5,2
100—299	24,0
300—499	31,4
500—1 000	32,4
Свыше 1 000	6,4
Итого	99,4

В этом случае нижняя граница каждой группы не должна совпадать с верхней границей предыдущей группы (такое совпадение возможно лишь в рядах с непрерывной вариацией признака).

Графики вариационных рядов

Вариационные ряды изображаются графически в виде линейных и плоскостных диаграмм в системе прямоугольных координат. По оси абсцисс откладываются значения варьирующего признака (варианты), а на оси ординат строится шкала для отсчета либо численности единиц (частот), либо их удельных весов в общем итоге (частостей), либо плотностей распределения.

Дискретные вариационные ряды, варианты которых выражены определенными числами, изображаются в виде линейных диаграмм.

Каждая частота или частость изображается точкой, абсцисса которой соответствует данному варианту, а ордината — частоте или частости. Эти точки последовательно соединяются отрезками, в результате чего получается ломаная кривая. Соединив крайние точки с осью абсцисс, получим замкнутый многоугольник — полигон распределения. Так, на рис. 4.7 изображен полигон, построенный по данным табл. 4.17.

Рис. 4.7. Распределение квартир по числу живущих в них лиц (полигон распределения)

Если значения признака выражены в виде интервалов вариации, то вариационный ряд изображается с помощью *гистограммы*. По оси абсцисс в соответствии с принятым масштабом откладываются границы интервалов, на которых строятся прямоугольники. Высоты последних пропорциональны плотностям распределения соответ-

ствующих интервалов. Плотность распределения — это число единиц совокупности, приходящееся на единицу ширины интервала. При равных интервалах плотности распределения пропорциональны частотам или частостям, которые и используются для построения прямоугольников. Изобразим, например, в виде гистограммы приводимые ниже данные о распределении населения СССР по возрасту.

Таблица 4.20

РАСПРЕДЕЛЕНИЕ НАСЕЛЕНИЯ СССР ПО ВОЗРАСТУ на 15 января 1970 г.)

	Число лиц (млн.)	Ширина возрастного интервала (лет)	Плотность распределения (гр. 1: гр. 2)
А	1	2	3
Менее 16 лет	74,7	16	4,7
16—44	104,9	29	3,6
45—59	33,3	15	2,2
60 лет и старше	28,5	40	0,7

Первый интервал содержит 16 однолетних возрастных групп, начиная с 0 лет (дети до 1 года) и кончая 15 годами. На каждую однолетнюю группу здесь приходится в среднем 4,7 млн. человек (74,7 : 16).

Плотности распределения второго и третьего интервалов равны 3,6 и 2,2 (млн. человек). Верхней границей последнего интервала условно будем считать 100 лет. Тогда его ширина составит 40 лет, а плотность распределения — около 0,7 (млн. человек на каждую однолетнюю возрастную группу).

Откладывая на оси абсцисс возраст, а на оси ординат — плотность распределения, получим гистограмму, изображенную на рис. 4.8. Общее число лиц в каждой возрастной группе на этом рисунке выражается площадью соответствующего прямоугольника.

Для графического изображения вариационного ряда используется также *кумулятивная кривая (кумулята)*. Для ее построения на оси абсцисс прямоугольной системы координат откладываются границы интервалов (или значения дискретного признака), а на оси ординат — нарастающие итоги частот, или частостей, соответствующие верхним границам интервалов. Пример построения кумуляты рассмотрен в гл. VI (см. рис. 6.1).

Если при группировке для каждой группы подсчитать не только число единиц, но и общий объем группировочного или какого-либо другого признака, то полученные данные будут характеризовать

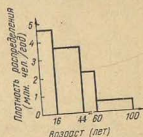


Рис. 4.8. Распределение населения СССР по возрасту (на 15 января 1970 г.)

Таблица 4.21

ГРУППИРОВКА ПРОМЫШЛЕННЫХ ПРЕДПРИЯТИЙ СССР ПО ОБЪЕМУ
ВАЛОВОЙ ПРОДУКЦИИ за 1972 г.

Группы предприятий по объему валовой продук- ции (млн. руб.)	В процентах к итогу			Нарастающие проценты к итогу		
	число пред- приятий	среднего- довая числен- ность рабоч- их	валовая продукция	число пред- приятий	среднего- довая числен- ность рабоч- их	валовая продукция
До 0,5	21,7	2,0	0,5	21,7	2,0	0,5
0,5—1,0	13,2	3,0	1,2	34,9	5,0	1,7
1,0—5,0	37,9	19,1	11,5	72,8	24,1	13,2
5,0—10,0	12,2	14,1	10,7	85,0	38,2	23,9
10,0—50,0	12,2	32,1	31,8	97,2	70,3	55,7
50,0—100,0	1,7	12,3	15,0	98,9	82,6	70,7
Более 100	1,1	17,4	29,3	100	100	100

распределение этого признака между единицами совокупности. На-
пример, из данных табл. 4.21 видно, что наиболее крупные по
объему продукции предприя-
тия, составляющие 1,1% обще-
го их числа, сосредоточивали
более 17% всех рабочих и да-
ли 29,3% всей валовой про-
дукции. В то же время не-
большие по объему продукции
предприятия, составляющие
21,7%, имели только 2,0% об-
щего числа рабочих и дали ме-
нее 1% всей продукции. Это
свидетельствует о концен-
трации рабочей силы и продукции
на крупных предприятиях.

Рис. 4.9. Концентрация рабочих и про-
изводства валовой продукции в про-
мышленности СССР в 1972 г. (кривые
концентрации)

числ., а нарастающие итоги других показателей — по оси ординат,
то получим ломаные кривые, характеризующие концентрацию
(в нашем примере — концентрацию рабочей силы и валовой про-
дукции). Эти кривые называются *кривыми концентрации (диффе-
ренции)*.

Для построения кривой концентрации рабочих нанесем на коор-
динатное поле точки, абсциссы которых равны нарастающим ито-

гам числа предприятий, а ординаты — нарастающим итогам числа
рабочих. Соединив последовательные точки, получим ломаную кри-
вую концентрации.

Если бы распределение рабочих между предприятиями было
равномерным, то кривая концентрации совпала бы с диагональю
от 0 до 100%, которая является графическим выражением равно-
мерного распределения и полного отсутствия концентрации. Чем
больше концентрация, тем более кривая отклоняется от диагонали.
Из рис. 4.9 видно, например, что концентрация валовой продукции
значительнее, чем концентрация рабочей силы.

6 ОРГАНИЗАЦИЯ И ТЕХНИКА СВОДКИ

Централизованная и децентрализованная сводка В соответствии с программой сводки, уста-
навливающей группировочные признаки,
число групп и макеты разрабаточных таб-
лиц, производится подсчет групповых и ито-
говых данных первичного статистического материала (сводка в уз-
ком понимании). По своей организации эта сводка может быть
централизованной и децентрализованной.

При *централизованной сводке* весь первичный материал, вклю-
чая первичные отчеты предприятий, сосредоточивается в одной
центральной организации, например ЦСУ СССР, где проходит раз-
работку от начала до конца. Именно так была организована раз-
работка материалов срочных переписей в период Великой Отече-
ственной войны, когда ЦСУ СССР запрашивало по телеграфу
необходимые сведения непосредственно у предприятий, минуя про-
межуточные инстанции государственной и ведомственной стати-
стики. Централизация сводки обеспечивает возможность широкой
механизации работ по подсчету данных, что ускоряет и удешев-
ляет сводку.

При *децентрализованной сводке* разработка материалов осу-
ществляется последовательными этапами по единой программе
разных организаций. Этим способом разрабатываются данные госу-
дарственной отчетности. Первичные материалы предприятий на-
чинают сдвигаться в пределах областей (краев) соответствующими
местными статистическими органами. Затем полученные резуль-
таты сдвигаются воедино ЦСУ союзных республик и, наконец, ЦСУ
СССР.

Точно так же разрабатываются материалы различных переписей
и учетов скота. При такой организации сводки местные органы
получают ее результаты в короткие сроки, облегчается проверка
актуальности представленных данных.

Вместе с тем ряд важнейших переписей, проводимых в Совет-
ском Союзе, например перепись населения, оборудования и дру-
гие, разрабатываются *комбинированно*: краткие основные итоги
подсчитываются в порядке децентрализованной сводки, начиная
от района и города, а полные — в порядке централизованной раз-
работки материала в ЦСУ СССР.

Техника машинизи- рованной сводки

По способу выполнения статистическая сводка может быть *машинизированной*, т. е. осуществляться с помощью счетно-аналитических либо электронных вычислительных машин, или *ручной*, т. е. выполняться вручную, без применения специальных машин.

Вычислительные перфорационные машины работают в комплексе, включающем перфораторы, контрольный, счетно-сортировочную машину, табулятор. Поэтому сводка на этих машинах предусматривает следующие операции: шифровку, перфорацию материала, проверку правильности перфорации, сортировку и табуляцию. Программа разработки обычно предполагает группировку, знаку присваивается отдельный шифр или условное обозначение, обычно в виде цифр или букв. К шифровке материала при машинизированной сводке предъявляются особые требования: шифры обязательно должны быть цифровые (одно- или многозначные в зависимости от числа групп по данному признаку), и при этом все без исключения группировочные признаки должны быть зашифрованы. Так, группировка населения по полу содержит всего две группы: мужчин и женщин. При ручной сводке можно обойтись и без шифровки материала по этому признаку, но при машинизированной сводке и этот признак должен быть зашиф-

рован. Вслед за шифровкой следует перфорация материала. Она состоит в том, что данные статистических формуляров переносятся на специальные карточки, которые называются перфорационными (перфокарты). Это строго стандартной формы (длина 187,4 мм, ширина 84,5 мм и толщина 0,18 мм) карточки из тонкого, но плотного картона. На их лицевой стороне отпечатана цифровая сетка в виде колонок цифр от нуля до девяти в каждой. Образующие эти цифры горизонтальные ряды называются *цифровыми позициями*. В СССР применяются 45- и 80-колонные перфокарты. Перенос зашифрованных числовых значений на перфокарты производится путем пробивки в них отверстий. Пробивка отверстий (*перфорация*) производится перфорационными машинами (перфораторами) по заранее установленной схеме (*макету перфокарты*), в которой для каждого признака отводится определенное число колонок. Например, при сводке материалов переписи населения принята таблица, в которой варианту «мужчины» отводится первая колонка, а варианту «женщины» — первая позиция; признаку «возраст» отводится вторая, третья и четвертая колонки и т. д. Тогда для мужчин в возрасте 35 лет будет составлена перфокарта, в которой будут пробиты отверстия: в первой колонке — нулевая позиция, во второй колонке (разряд сотен) — нулевая позиция, в третьей колонке — третья позиция и в четвертой колонке — пятая позиция (см. рис. 4.10).

На каждую единицу наблюдения составляется, как правило, отдельная перфокарта.

Правильное заполнение перфорационных карточек имеет очень важное значение для последующих этапов сводки. Поэтому все перфокарты подвергаются проверке при помощи специальных контрольных машин (контрольников). Для этого отперфорированные карты одна за другой вводятся под воспринимающее устройство контрольника, а на его цифровой клавиатуре набирают те данные, которые должны быть «записаны» на перфокарте. Если при нажатии клавиши пробивка на перфокарте совпадает с набором на клавиатуре, то перфокарта передвигается в следующую колонку, если же такого совпадения нет, то перфокарта задерживается на контролируемой колонке, что свидетельствует о наличии ошибки. Перфокарта, в которой обнаружена ошибка, перебивается заново.

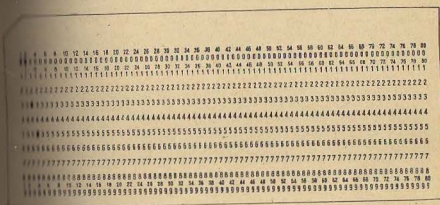


Рис. 4.10. Перфорационная карта

Отперфорированные и проверенные перфокарты поступают в сортировочную машину, которая производит раскладку перфокарт по установленным группировочным признакам и подсчет как общего числа пропущенных через машину карточек, так и количества карточек в каждой отдельной группе. Перфокарты движутся по конвейеру и в зависимости от того, на какой позиции в данной колонке пробито отверстие, сбрасываются в соответствующий карман сортировки. При сбрасывании карточка задает светчик, который автоматически подсчитывает число карточек, попавших в тот или иной карман.

Разгруппированные по одному признаку карточки подвергаются затем сортировке по второму, третьему и т. д. признакам в зависимости от программы сводки.

Если сводка должна дать только число единиц по группам, то на этой стадии машинизированная сводка заканчивается. Показания счетчиков записываются в соответствующие таблицы (группные итоги), а общие итоги подсчитываются с помощью суммирующих машин. Если же сводка должна также дать и значения количественных признаков (например, оборот торгующих

организаций), то перфокарты после сортировки поступают на *табуляцию*. Табулятор — это машина, которая подсчитывает количество карт (единиц) и объемы количественных признаков по отдельным группам и в целом, а результаты подсчетов печатает на ленте, называемой *табуляграммой*. Данные табуляграмм заносятся в статистические таблицы.

Сводка статистического материала на машинах обходится значительно дешевле ручной. Вместе с тем она резко повышает производительность труда, ускоряет процесс разработки и гарантирует при условии правильной перфорации материала от ошибок, которые могут быть при ручной сводке. Эти преимущества машинной сводки особенно наглядны при разработке материалов переписей населения. Ручная разработка материалов переписи населения такой страны, как Советский Союз, потребовала бы многих лет, а использование комплекса высокопроизводительных вычислительных машин позволило получить первые итоги переписи населения СССР 1970 г. менее чем через 4 месяца после ее завершения.

Новые задачи статистики, вытекающие из Программы КПСС, требуют от статистической науки дальнейшего совершенствования методов сводки и обработки материалов статистического наблюдения. Статистическая информация в объеме, в котором она поступает в органы государственной статистики, уже в настоящее время может быть переработана лишь с помощью машинной техники.

Внедрение электронных вычислительных машин (ЭВМ), выполняющих сотни тысяч операций в секунду, связано с коренной перестройкой организации сбора и сводки статистических материалов.

Паряду с получением информации в виде текстов или иных форм статистической отчетности начинается передача экономической информации с рабочих мест, где первичные данные с помощью различных счетчиков и датчиков автоматически передаются непосредственно в обрабатывающую систему. Этим достигается высокая достоверность исходных данных, громадное ускорение передачи информации и ее обработки.

Меняются и способы передачи сводной информации в вычислительные центры. Вместо почтовых отправок на первое место выдвигается телеграфный способ передачи информации, в частности абонентский телеграфный способ передачи информации. При этом способе экономится время прохождения и обработки информации на промежуточных этапах и создается возможность агрегировать стартовые телеграфные аппараты с устройствами для автоматизации занесения принимаемой информации на перфокарты.

Для более полного использования возможностей ЭВМ перфокарты как технические носители информации заменяются перфолентами, магнитными картами и лентами, микропленкой, что обеспечивает значительное сокращение затрат труда по переносу на них первичной информации и быстрый ввод данных в электронную

вычислительную машину. Разработанный научно-исследовательской лабораторией ИСУ СССР метод микрофильмирования статистических документов с непосредственным вводом информации в ЭВМ значительно упрощает процесс шифровки по сравнению с ручной перфорацией, сокращает трудоемкость процесса почти в 10 раз, исключает ошибки при переносе данных на перфокарту и обеспечивает полное использование быстродействия ЭВМ.

Мощная вычислительная техника может найти полное применение только в условиях социалистического общества. Здесь имеются возможности комплексной автоматизации обработки экономической информации для получения экономических показателей работы отдельных предприятий, групп предприятий в пределах экономического района, всего экономического района, республики, наконец, всего народного хозяйства.

Техника ручной сводки

Ручная сводка применяется обычно при небольшом объеме материала и может осуществляться двумя способами: подсчеты цифровых данных либо с карточек, либо со списков.

При ручной сводке с карточек необходимо выполнить следующие операции: разметку (шифровку), раскладку материала по группам, подсчет числа карточек, входящих в каждую группу, подсчет итоговых данных по признакам, включенным в разметку, и, наконец, занесение результатов сводки в табличные формы.

Для осуществления принятой в плане группировки карточки должны быть разложены по предусмотренным для каждого разрабатываемого признака группам. Если варианты признака легко запоминаются (например, при группировке населения по полу, партийной принадлежности и т. д.), то раскладка карточек производится непосредственно. Если же варианты признака сложны (например, занятия населения, национальность и др.), то предварительно производится шифровка материала. Шифр обычно проставляется в карточке на специально отводимом для этого месте.

После того как карточки зашифрованы, производится их раскладка. Как правило, данные статистического наблюдения сводятся по ряду признаков, взятых нередко в комбинации. Поэтому нужно установить такой порядок раскладки карточек, чтобы они подвергались наименьшему числу перемещений. Для этого один из признаков, являющийся общим для нескольких группировок, должен быть принят за постоянный. Весь материал следует расположить по этому признаку и оставить в этой раскладке до тех пор, пока данный признак будет встречаться в комбинации с какими-либо другими признаками. Когда признак, по которому сделана раскладка, полностью исчерпан (во всех сочетаниях с другими признаками), должна быть сделана новая раскладка опять-таки по такому признаку, группировка по которому встречается чаще других.

Поясним это таким примером. Допустим, что материалы обследования рабочих завода нужно обработать так, чтобы получить комбинационные таблицы, характеризующие распределение рабочих: а) по общему стажу работы и по квалификации; б) по размеру заработной платы и по квалификации; в) по возрасту и по квалификации. Чтобы наиболее рационально выполнить задачу сводки, целесообразно раньше всего произвести раскладку карточек на группы по квалификации (разрядам), а затем внутри каждой квалификационной группы производить раскладку по другим признакам.

После того как раскладка карточек закончена, производится их подсчет. Если требуется выяснить только численность единиц данной группы, то подсчитывают число карточек, отнесенных в ту или иную группу. Но часто бывает необходимо наряду с численностью единиц группы подсчитать и суммы значений отдельных признаков у этих единиц.

Для облегчения подсчетов цифровые данные записываются на краях карточек. Накладывая карточки друг на друга таким образом, чтобы край с цифрами внизу лежащей карточки не закрывался карточкой, лежащей сверху, можно быстро подсчитать итоги по каждому из показателей. Результаты подсчетов заносятся в соответствующие клетки сводной таблицы. Заключительной стадией сводки является подсчет итогов по строкам и графам.

При *ручной сводке со списков* различают два случая: 1) первичный материал занесен в списки в том порядке, в каком он должен сводиться, и 2) первичный материал, внесенный в списки, нужно группировать в ином порядке.

В первом случае техника подсчета сводится к простому подведению итогов по отдельным показателям. Во втором случае (особенно при большом объеме материала) целесообразно по каждой строке списка составлять специальные фишки (карточки, в которые заносятся значения всех показателей в таком порядке, в каком они следуют в списке). Техника сводки с фишек аналогична технике сводки с карточек.

Если объем материала невелик, подсчет численности каждой группы возможен с помощью вспомогательных таблиц с такими же заголовками граф и строк, как и в таблицах, в которые будут внесены итоги сводки, но с более широкими клетками. Каждое единичное показание из списка заносится в соответствующую клетку вспомогательной таблицы либо в виде штриха (при подсчете только числа единиц), либо в виде числового значения (при подсчете суммы значений признака).

Глава V

АБСОЛЮТНЫЕ И ОТНОСИТЕЛЬНЫЕ ВЕЛИЧИНЫ

1. АБСОЛЮТНЫЕ ВЕЛИЧИНЫ

Виды и способы подсчета абсолютных величин

Результаты статистического наблюдения и сводки его материалов выражаются прежде всего в абсолютных величинах (показателях).

Абсолютные величины показывают размеры (уровни, объемы) общественных явлений в данных условиях места и времени или величину признаков, характеризующих эти явления, например: объем валовой продукции завода в данном году, численность скота в колхозе на определенную дату, количество тонн груза, перевезенного железнодорожным транспортом в отчетном периоде, величина заработной платы данного работника, цена одного килограмма масла на колхозном рынке и т. д.

Абсолютные величины подразделяются на индивидуальные и суммарные (т. е. итоговые).

Индивидуальными называются абсолютные величины, выражающие размеры количественных признаков у отдельных единиц той или иной совокупности объектов (например, величина заработной платы отдельного рабочего, размер посевной площади пшеницы в отдельном совхозе и т. п.). Они получают непосредственно в процессе статистического наблюдения и фиксируются в первичных учетных документах. Индивидуальные абсолютные величины служат основой любого статистического исследования.

В отличие от индивидуальных суммарные абсолютные величины характеризуют итоговую величину признака по определенной совокупности объектов, охваченных статистическим наблюдением. Они получаются либо путем прямого подсчета числа единиц наблюдения, либо в результате суммирования значений признака у отдельных единиц совокупности. Так, в результате подсчета зарегистрированных в процессе переписи населения лиц получают итоговые абсолютные данные о численности населения по отдельным областям, республикам и в целом по СССР; о численности мужчин и женщин и т. п. В результате сводки годовых отчетов колхозов получают абсолютные итоговые данные о ко-

личестве колхозов, о количестве полученной продукции и другие суммарные показатели.

В ряде случаев суммарные абсолютные величины получаются в результате сводки данных статистического наблюдения, а путем специальных расчетов. При помощи таких расчетов определяется, например, перспективная численность населения, количество сельскохозяйственной продукции, произведенной в подсобных хозяйствах населения отдельных областей (так, валовой сбор картофеля исчисляется путем умножения посевной площади картофеля на среднюю его урожайность с гектара и т. п.). Практическая потребность в подобных расчетах обычно обуславливается отсутствием соответствующих данных в первичных учетно-статистических материалах. Особенно широко применяются расчетные абсолютные величины в планировании.

Абсолютные величины имеют большое практическое и познавательное значение. В абсолютных величинах выражаются размеры всех видов национального богатства страны, наличие и движение материальных ресурсов и денежных средств на различных участках коммунистического строительства. В абсолютных величинах устанавливается большинство плановых заданий по развитию народного хозяйства и культуры и осуществляется контроль за их выполнением. Абсолютные величины необходимы для хозяйственных расчетов и широко используются в статистико-экономическом анализе. Они служат исходными данными для всех форм и приемов количественной характеристики общественных явлений.

Абсолютные величины представляют собой именованные числа. Они характеризуют размеры качественно определенных явлений и выражаются в определенных единицах измерения. Этим они отличаются от отвлеченных математических величин.

В зависимости от сущности изучаемого явления и конкретных задач исследования применяются натуральные и денежные (ценностные) единицы измерения.

Натуральными называются такие единицы измерения, которые соответствуют естественным (физическим) свойствам данного предмета и выражаются в мерах длины, площади, объема, веса и т. п. или количеством единиц (штук), числом фактов и событий. Так, единиц измерения численности населения является один человек, сбор зерна измеряется в тоннах, центнерах, а иногда в пудах, выпуск тканей — в погонных и квадратных метрах и т. п. Используются в статистике и единицы измерения времени. Так, возраст, стаж работы измеряются в годах, в месяцах.

В некоторых случаях использование одной единицы измерения не дает полной характеристики размера явления и тогда приходится пользоваться двумя единицами измерения. Так, например, тракторы по их количеству можно учитывать в так называемых физических единицах. Однако отдельные тракторы могут иметь различную мощность, а следовательно, и различную потребитель-

скую стоимость. Поэтому для правильного представления о производстве тракторного завода нужно учесть выпуск тракторов в двух различных измерениях: в штуках и по общей (суммарной) мощности. По этим же соображениям ткани учитываются в погонных и квадратных метрах и т. д.

Иногда применяются комбинированные единицы измерения. Так, например, грузооборот измеряется в тонно-километрах и тонно-милях, электроэнергия — в киловатт-часах и сидо-часах, затраты труда — в человеко-часах, человеко-днях и т. п.

Данные о количестве различных продуктов, выраженные в натуральных единицах измерения, не допускают суммирования. Поэтому в статистике с целью получения общего итога близких по своему потребительскому назначению продуктов иногда используются условные натуральные единицы измерения. Для этого прежде всего находятся так называемые переводные коэффициенты, выражающие соотношения между натуральными единицами измерения различных продуктов по какому-либо признаку (потребительным свойствам, трудоемкости, себестоимости и т. д.). Затем по найденным коэффициентам эти продукты пересчитываются на один продукт, натуральная единица которого принята в качестве условной единицы измерения.

Допустим, что мыловаренный завод выпустил 5 тыс. т мыла с 40%-ным содержанием жирных кислот, 3 тыс. т — с 60%-ным содержанием жирных кислот и 2 тыс. т — с 80%-ным содержанием жирных кислот. Тогда вся продукция завода в пересчете на условные 40%-ное мыло составит: $5 \times 1,0 + 3 \times 1,5 + 2 \times 2,0 = 13,5$ тыс. т (переводный коэффициент для 40%-ного мыла равен 1,0, для 60%-ного — $60:40 = 1,5$ и для 80%-ного — $80:40 = 2,0$). Таким же способом производится пересчет различных видов топлива — в условные 7000-калорийное топливо и т. д.

Условные натуральные единицы измерения имеют довольно ограниченное применение в практике и совершенно непригодны для исчисления общего объема разнородной продукции.

Мерой общего объема разнородных материальных благ служат обычно стоимостные (денежные) единицы измерения (рубль, тысяча рублей и т. д.). Экономический оборот общественного продукта происходит в денежной форме, в связи с чем стоимостные единицы измерения имеют широкое практическое применение.

3. ОТНОСИТЕЛЬНЫЕ ВЕЛИЧИНЫ

Сущность и значение относительных величин

Наряду с абсолютными величинами большое значение в статистике имеют относительные величины.

Абсолютные величины, взятые сами по себе, зачастую не дают надлежащего представления об изучаемых явлениях и процессах. Известно, например, что в 1973 г. в СССР было произведено 1345 млн. кв. м шелковых тканей. Для того чтобы

судить о том, много это или мало, достаточно для удовлетворения потребностей или нет, чтобы надлежащим образом оценить приведенную величину, нужно сравнить ее с другими величинами.

В нашем примере, сопоставив производство шелковых тканей в 1973 г. с производством их в 1972 г. (1270 млн. кв. м) и в 1971 г. (1190 млн. кв. м), мы обнаружим, что в нашей стране происходит неуклонный рост производства шелковых тканей. Сопоставив годовой выпуск шелковых тканей с численностью населения СССР, можно определить также размер их производства на душу населения.

Сравнение является основным приемом оценки статистических данных и составной частью всех методов их анализа. Однако простое сопоставление двух величин, как это было сделано выше, недостаточно для точной оценки их соотношения. Это соотношение нужно измерить. Роль меры соотношения — выполнять относительные величины.

В нашем случае можно, например, определить, что в 1973 г. производство шелковых тканей в СССР увеличилось по сравнению с 1971 г. более чем в 1,1 раза (1345 : 1190), а по сравнению с 1972 г. — на 5,9% (1345 : 1270). Разделив объем производства шелковых тканей в 1973 г. на среднюю численность населения за этот год (249,7 млн. человек), найдем, что на душу населения было произведено более 5 кв. м шелковых тканей (1345 : 249,7).

В отличие от абсолютных статистических показателей относительные величины представляют собой производные величины. Они получаются не в результате простого суммирования, а путем относительного (кратного) сравнения между собой абсолютных величин и выражают, таким образом, количественные соотношения, присущие конкретным общественным явлениям и процессам.

Относительные величины исчисляются для самых разнообразных целей. С их помощью выражаются соотношение производства средств производства и предметов потребления, структура посевных площадей, темпы роста производства и другие количественные характеристики явлений общественной жизни, позволяющие глубже проникнуть в существо этих явлений, составить о них более полное и квалифицированное суждение.

Формы выражения и виды относительных величин

В зависимости от характера изучаемого явления и конкретных задач исследования относительные величины могут иметь различную форму (внешний вид) выражения. Наиболее простой формой выражения относительной величины является число (целое или дробное), показывающее, во сколько раз одна величина больше другой, принятой за базу сравнения, или какую часть ее составляет. Так, если известно, что на 1 января 1974 г. население г. Москвы составляло 7 млн. 528 тыс. человек, а г. Одессы — 981 тыс. человек, то отношение численности населения Москвы к численности населения Одессы можно выразить так: $7528 : 981$

или 7,7. Число 7,7 является относительной величиной, показывающей, что численность населения Москвы в 7,7 раза больше, чем Одессы. За базу сравнения можно принять население Москвы.

Такая отношение $981 : 7528 = \frac{1}{7,7}$ покажет, что население Одессы составляет восьмую часть населения Москвы.

Другой (часто более удобной) формой относительных чисел являются широко распространенные в практике процентные отношения, при которых базисная величина принимается за 100. Основное преимущество процентных отношений заключается в том, что они легко воспринимаются.

В ряде случаев базисная величина принимается не за 100%, а за 1000 промилле (‰). Так, соотношение численности студентов высших учебных заведений СССР в 1973/1974 учебном году (4,67 млн. человек) и общей численности населения СССР в начале 1974 г. (250,9 млн. человек) составляет $(4,67 : 250,9) \times 1000 = 1,86\%$, или $(4,67 : 250,9) \times 1000 = 18,6\text{‰}$. Это значит, что из каждых ста человек населения обучалось в вузах около 2 (1,9) человек, или из каждой тысячи — почти 19 человек. Более удобна здесь вторая форма выражения. Чтобы избежать трудновоспринимаемых дробных относительных величин, базисная величина принимается иногда за 10 000, за 100 000 и т. д.

В статистической практике применяются все указанные формы выражения относительных величин. Однако в каждом отдельном случае следует выбирать такую из этих форм, которая легче воспринимается и с наибольшей наглядностью выражает искомое соотношение.

Относительные величины различаются не только по форме выражения, но и по сущности выражаемых ими количественных соотношений. По этому признаку относительные величины можно подразделить на следующие виды: динамики, выполнения плана, структуры, интенсивности, координации и сравнения:

Относительными величинами динамики — темпами роста — называются показатели, характеризующие изменения уровня отдельных явлений во времени.

Они исчисляются как отношение уровней данного явления, относящихся к разным периодам или моментам времени, и выражаются в виде процентов или коэффициентов. Если, например, в 1973 г. завод выпустил 500 станков, а в 1974 г. — 630 станков, то относительный показатель, характеризующий динамику (темп роста), составит $126\% [(630 : 500) \times 100]$. Он показывает, что выпуск станков в 1974 г. составил 126% выпуска 1973 г., т. е. увеличился за год на 26%.

Относительные величины динамики применяются в различных отраслях советской статистики. Очень часто эти показатели встречаются в докладах и статьях на политико-экономические темы. Широко используются они и в анализе хозяйственной деятельности отдельных предприятий и организаций.

Наряду с фактическими темпами роста часто исчисляются также плановые темпы роста, которые выражают величину конкретных плановых заданий, установленных на тот или иной период, по отношению к фактически достигнутому ранее уровню.

В перспективных, годовых, квартальных и других народнохозяйственных планах задания в процентах к базисному уровню, т. е. плановые темпы роста, даются по ряду важнейших показателей. Так, в государственном плане развития народного хозяйства на 1975 г. предусматривается увеличение валовой продукции промышленности на 47% по сравнению с 1970 г., повышение производительности труда в промышленности — на 38,8% и в строительстве — на 37%.

Плановые темпы роста рассчитываются в виде отношения:

$$\frac{\text{Плановое задание на предстоящий период}}{\text{Фактическое выполнение за базисный период}}$$

Это отношение показывает, во сколько раз плановое задание больше уровня, достигнутого в базисном периоде (или какую часть его составляет). Для того чтобы выразить плановые задания в процентах, достаточно данное отношение умножить на 100. Если, например, в 1973 г. завод выпустил 500 станков, а по плану на 1974 г. предусматривается выпустить 600, то плановый темп роста составит: $600 : 500 = 1,2$, или 120%. Это означает, что планируется увеличение выпуска станков в 1,2 раза, или на 20%.

Относительные величины выполнения плана

Другим видом относительных показателей являются относительные величины выполнения плана. Их применение обусловлено плановой экономикой.

Относительными величинами выполнения плана называются показатели, характеризующие степень выполнения планового задания за данный период времени.

Относительные величины выполнения плана имеют огромное значение в статистике. С их помощью осуществляется контроль за ходом выполнения государственного плана развития народного хозяйства во всех его звеньях. Данные о выполнении годовых и квартальных народнохозяйственных планов по стране в целом, республикам и областям систематически публикуются в печати и служат целям оперативного руководства хозяйством. Они помогают определить передовые и отстающие участки (предприятия, отрасли, министерства, отдельные районы, области) и своевременно принять необходимые меры.

Относительные величины выполнения плана исчисляются обычно в виде процентного отношения:

$$\frac{\text{Фактическое выполнение}}{\text{Плановое задание}} \times 100,$$

где фактическое выполнение и плановое задание берутся за один и тот же период или на один и тот же момент времени.

Вали, например, в 1974 г. завод выпустил 630 станков, а планировалось выпустить 600 станков, то процент выполнения плана составит $105 [(630 : 600) \times 100]$. Это означает, что план выпуска станков был выполнен на 105%, т. е. перевыполнен на 5%.

Аналогично вычисляется процент выполнения плана и в тех случаях, когда само плановое задание выражено в процентах к базисному уровню. Разница заключается лишь в том, что и в числительная величина, представляющая собой процентное отношение фактического выполнения к базисному уровню, т. е. фактический темп роста.

Если планируемый показатель *снижается* при улучшении работы предприятия (например, себестоимость продукции), то *превышение* его фактических размеров над плановыми означает *невыполнение*, и, наоборот, *снижение* его фактической величины на сравнении с планом свидетельствует о *перевыполнении* плана. Поэтому и процент выполнения плана, рассчитанный указанным выше способом, оказывается тем меньше, чем лучше выполнен план, и наоборот.

Допустим, что производственным планом предприятия намечалось снижение себестоимости продукции за год на 10%, а фактически она снизилась только на 4%. Следовательно, план по снижению себестоимости продукции не выполнен. Однако показатель выполнения плана, исчисленный как отношение отчетной величины к плановой, оказывается более 100%:

$$\frac{100 - 4}{100 - 10} \times 100 = \frac{96}{90} \times 100 = 106,7\%.$$

Полученный результат показывает, что фактическая себестоимость выше плановой на 6,7%, и означает поэтому не перевыполнение плана снижения себестоимости продукции, а его недовыполнение на 6,7% ($106,7 - 100$).

Нетрудно убедиться, что относительная величина выполнения плана связана с фактическим и плановым темпами роста. Если обозначить абсолютные величины базисного и отчетного периодов соответственно через y_0 и y_1 , а абсолютную величину планового задания через $y_{пл}$, то будем иметь:

$$\frac{y_{пл}}{y_0} \cdot \frac{y_1}{y_{пл}} = \frac{y_1}{y_0} \quad \frac{y_{пл}}{y_0} \cdot \frac{y_1}{y_{пл}} = \frac{y_1}{y_0} \quad (5.1)$$

Таким образом, произведение планового темпа роста и степени выполнения плана равняется фактическому темпу роста по отношению к той же базе, которая принята в плане. Для иллюстрации указанной взаимосвязи вычислим фактический темп роста, используя приведенные выше расчеты планового темпа роста и степени выполнения плана по выпуску станков: $1,2 \times 1,05 = 1,26$, или 126%.

Взаимосвязь указанных трех относительных величин может быть использована для исчисления по двум известным величинам

неизвестной третьей величины. Известно, что валовая продукция промышленности СССР в 1973 г. увеличилась по сравнению с 1972 г. на 7,4% вместо 5,8% предусмотренных планом. Следовательно, $(y_{73} : y_{72}) \times 100 = 107,4$ и $(U_{пл\ 73} : U_{пл\ 72}) \times 100 = 105,8$. Отсюда выполнение плана в 1973 г. составляет:

$$\left(\frac{U_{73}}{U_{72}} : \frac{U_{пл\ 73}}{U_{пл\ 72}} \right) \times 100 = \frac{U_{73}}{U_{пл\ 73}} \times 100 = (107,4 : 105,8) \times 100 = 101,5\%$$

Относительные величины структуры

Относительными величинами структуры называются показатели, характеризующие доли (удельные веса) составных частей целого в их общем итоге. Например, удельный вес мужчин и удельный вес женщин в общей численности населения, удельные веса производства того или иного продукта колхозами, совхозами и т. д. в общем размере производства данного продукта.

При вычислении относительных показателей структуры целое принимаем за базу (основание) сравнения и находим долевое или (чаще всего) процентное отношение отдельных частей к целому, т. е. к их общему итогу.

В качестве примера приведем возрастную структуру населения СССР.

Таблица 6.1

ВОЗРАСТНАЯ СТРУКТУРА НАСЕЛЕНИЯ СССР
по данным переписей населения)

	1959 г.		1970 г.		Темп роста в процентах (гр. 3 : гр. 1 × 100)
	Тысяч человек	В процентах к итогу	Тысяч человек	В процентах к итогу	
А	1	2	3	4	5
Все население	190 678	100	241 720	100	126
в том числе:					
0—9 лет	43 476	22,8	44 986	18,6	103
10—19	41 395	21,7	46 987	19,5	113
20—29	34 306	18,0	30 875	12,8	90
30—39	28 556	15,0	37 739	15,6	132
40—49	17 379	9,1	31 259	13,0	179
50—59	12 533	6,6	21 091	8,7	168
60—79	12 090	6,3	25 619	10,6	212
80 лет и старше	907	0,5	2 895	1,2	319

В гр. 2 и 4 таблицы показаны относительные величины структуры, выражающие удельные веса отдельных возрастных групп населения в его общей численности.

Относительные величины, характеризующие внутреннюю структуру общественных явлений, широко применяются в различных отраслях статистики и в экономическом анализе. При их помощи устанавливается классовая, возрастная и национальная структура

населения, структура общественного продукта и основных фондов народного хозяйства, структура товарооборота и т. п.

Относительные величины структуры используются и для выявления различий в составе двух или нескольких совокупностей по какому-либо общему для них признаку. Примером этого могут служить данные следующей таблицы.

Таблица 5.2

СТРУКТУРА ОСНОВНЫХ ФОНДОВ ХОЗЯЙСТВЕННЫХ ПРЕДПРИЯТИЙ И ОРГАНИЗАЦИЙ ПРОМЫШЛЕННОСТИ И СЕЛЬСКОГО ХОЗЯЙСТВА СССР на 1 ЯНВАРЯ 1972 г.
(на восстановительной стоимости; в процентах)

Элементы основных фондов	Промышленность	Сельское хозяйство
Все основные фонды	100	100
в том числе:		
здания	30,3	41,7
сооружения	19,5	12,0
предметные устройства	11,6	2,2
машины и оборудование	35,8	19,0
транспортные средства	2,1	4,1
инструмент, производственный и хозяйственный инвентарь	0,6	1,4
прочие основные фонды	0,1	19,6

Из приведенной таблицы видно, что в структуре основных фондов этих двух отраслей народного хозяйства имеются существенные различия: в промышленности высок удельный вес машин и оборудования, а в сельском хозяйстве наибольший удельный вес занимают здания.

Сопоставление структуры той или иной совокупности за два или несколько последовательных периодов позволяет установить структурные изменения, происшедшие в ее составе, их направление и тенденцию. Так, сопоставляя возрастную структуру населения СССР за 1939 и 1970 гг. (см. табл. 5.1), можно видеть, как изменились за этот период численность и удельные веса каждой возрастной группы населения.

На этом же примере видна взаимосвязь показателей структуры и показателей динамики. Эта взаимосвязь выражается в зависимости изменения структуры (так называемых структурных сдвигов) от неравномерности темпов роста составных частей целого. Зависимость заключается в том, что *удельные веса тех частей, которые растут быстрее, чем целое, повышаются, а удельные веса частей, растущих менее высокими темпами, чем целое, снижаются.*

Данные граф 2, 4 и 5 табл. 5.1 могут служить наглядной иллюстрацией этого. Так, темп роста численности населения в каждой из пяти возрастных групп старше 30 лет (132%, 179%, 168%, 312% и 319%) был выше темпа роста численности всего население,

Таблица 5.3

ПРОИЗВОДСТВО ВАЖНЕЙШИХ ВИДОВ ПРОДУКЦИИ НА ДУШУ НАСЕЛЕНИЯ
В СССР И США

Наименование продукции	СССР			США
	1913 г.	1940 г.	1973 г.	1973 г.
Хлеб — кг	27	94	526	661
Мясо — кг	65	159	1718	2154
Электроэнергия — кВт·ч	13	255	3662	9805
Удобрения — кг	11	30	430	372
Искусственные удобрения — кг усл.	0,6	17	290	368
Шерсть козальная — пар	4	1,1	2,7	2,3
Запас несок (из отечественного сырья) — кг	8,6	11,1	33,8	24
Молоко животное — кг	2,3	1,9	5,4	2,0
Молочко — кг	185	172	354	249

ния (126%), в результате чего удельный вес этих групп в 1970 г. по сравнению с 1939 г. повысился. Численность населения первых трех возрастных групп (до 30 лет) росла медленнее, чем все население (103%, 113% и 90%), вследствие чего их удельный вес в 1970 г. оказался ниже, чем в 1939 г.

Относительные величины интенсивности называются показателями, характеризующими степень распространения или развития данного явления в определенной среде. Они исчисляются как отношение абсолютной величины данного явления к размеру среды, в которой оно развивается.

Примером относительных величин интенсивности могут служить коэффициенты рождаемости, смертности и естественного прироста населения, используемые при анализе воспроизводства населения. Абсолютные числа родившихся и умерших, а также превышение числа родившихся над числом умерших, т. е. естественный прирост, при прочих равных условиях тем больше, чем больше численность населения. Поэтому для измерения интенсивности процессов рождаемости, смертности и естественного прироста населения соответствующие абсолютные числа относят к средней численности населения. Для более удобного восприятия эти коэффициенты выражаются в промилле, т. е. рассчитываются на 1000 человек населения. Так, в СССР в 1973 г. коэффициент рождаемости составил 17,6‰, коэффициент смертности — 8,7‰, коэффициент естественного прироста — 8,9‰. Следовательно, в расчете на каждую тысячу человек населения в 1973 г. родилось около 18 детей, умерло 8,7 человека, а прирост составил 9 человек.

Можно привести много других примеров, иллюстрирующих значение и необходимость применения этого вида величин в конкретном анализе. С их помощью определяется, например, плотность населения (число жителей в среднем на 1 кв. км), обеспеченность населения врачебной помощью (число врачей в расчете на 10000 человек населения), уровень вооруженности труда основными фондами (стоимость основных фондов в расчете на одного рабочего в наиболее заполненной смене) и т. п.

К относительным величинам интенсивности, как особой их разновидности, относятся также показатели, характеризующие интенсивность процессов производства и потребления материальных благ. Ярким примером относительных величин интенсивности такого вида могут служить показатели производства продукции на душу населения.

Эти показатели представляют собой отношение количества той или иной продукции к средней численности населения за соответствующий год. Они характеризуют уровень экономического развития страны.

Из таблицы видно, какой значительный шаг сделала наша страна в этом отношении и какие рубежи нужно еще преодолеть для того, чтобы догнать и превзойти США в этой области, как это предусматривается Программой КИСС.

Примерами относительных величин, характеризующих интенсивность процессов в сфере производства и потребления, являются и другие статистические показатели, широко применяемые в экономическом анализе, например производство национального дохода на душу населения, уровень потребления, исчисляемый как отношение количества потребленных продуктов к среднегодовой численности населения, производство сельскохозяйственных продуктов на 100 га земельных угодий, выход продукции на 1 рубль основных производственных фондов, коэффициенты энерго- и электрооборуженности труда и т. п.

В отличие от других видов относительных величин показатели интенсивности являются обычно именованными числами и имеют размерность тех абсолютных величин, соотношением которых они выражаются.

При исчислении показателей интенсивности большое значение имеет правильный выбор базы, с которой следует сопоставлять изучаемое явление. За базу сравнения нужно принимать, как правило, только ту совокупность (среду), в которой может иметь место изучаемое явление. Например, выход молока и мяса крупного рогатого скота рассчитывается на 100 га всех сельскохозяйственных угодий, а выход свиного мяса — на 100 га пашни.

Относительные величины координации или наглядности характеризуют отношения отдельных частей совокупности к одной из них, взятой за базу сравнения. Такими показателями являются, например, число инженерно-технических работников и число служащих на 100 рабочих, количество гектаров технических (или кормовых) культур на 100 гектаров зерновых культур, число женщин на 100 мужчин и т. п. Характеризуя соотношения между важнейшими составными частями целого, относительные величины данного вида позволяют контролировать соблюдение необходимых пропорций между ними.

Относительные величины координации представляют собой отношение абсолютных величин, имеющих одинаковую единицу измерения, и выражаются потому в процентах, промилле или кратных отношениях, а не в именованных числах.

Относительные величины сравнения называются относительными показателями, характеризующие сравнительные размеры одноименных величин, относящихся к одному и тому же периоду или моменту времени, но к различным объектам или территориям. Обычно они исчисляются в процентах или в кратных отношениях, показывающих, во сколько раз одна сравниваемая величина больше (меньше) другой.

Примером относительных величин сравнения могут служить следующие данные о соотношении производства некоторых видов продукции в СССР и США в 1973 г.

Таблица 54

**ПРОИЗВОДСТВО ВАЖНЕЙШИХ ВИДОВ ПРОМЫШЛЕННОЙ ПРОДУКЦИИ
в СССР и США в 1973 г.**

	СССР	США	СССР в процентах к США
Сталь — млн. т	131	139	94
Нефть — млн. т	429	453	95
Уголь (товарный) — млн. т	615	542	113
Цемент — млн. т	110	78	141
Электроэнергия — млрд. кВт·ч	915	2 068	44

В последней графе таблицы приведены относительные величины сравнения, выраженные в процентах. Они показывают, что по производству стали, нефти и электроэнергии СССР пока еще отстает от США, а по производству угля и цемента — опережает США и вышел на первое место в мире.

Относительные величины сравнения широко применяются при сравнительной оценке показателей работы отдельных предприятий, городов, районов, областей и т. д.

3. ОБЩИЕ ПРИНЦИПЫ ПОСТРОЕНИЯ И ИСПОЛЬЗОВАНИЯ АБСОЛЮТНЫХ И ОТНОСИТЕЛЬНЫХ СТАТИСТИЧЕСКИХ ВЕЛИЧИН

Правильное применение абсолютных и относительных величин для характеристики общественных явлений и процессов возможно только на основе соблюдения некоторых общих принципов.

Важнейшим и необходимым условием правильной количественной характеристики общественных явлений и процессов является понимание их сущности, специфических особенностей и законов их развития. Поэтому при расчете и использовании абсолютных и относительных величин необходимо прежде всего учитывать

специфику, конкретные особенности и условия развития изучаемых явлений и процессов.

Общественные явления отличаются многообразием и сложностью. Их размер и количественные соотношения могут изменяться в зависимости от времени и места с различной быстротой и в неодинаковых направлениях. Отсюда вытекает необходимость дифференцированного подхода к изучению конкретных общественных явлений, а следовательно, и дифференцированного использования в экономическом анализе абсолютных и относительных величин. Без этого невозможно было бы выявить особенности развития тех или иных общественных явлений, определить передовые и отстающие участки. Например, при выполнении плана по отрасли в целом могут быть отдельные предприятия, не выполнявшие план и перевыполнившие его. Только дифференцированный подход к изучению указанных явлений позволяет выявить все их особенности.

Важным условием правильного применения абсолютных и относительных величин является необходимость их комплексного использования в статистическом исследовании. Это условие вытекает непосредственно из характера взаимосвязи абсолютных и относительных величин. Взаимосвязь между ними ясна. Относительные величины являются производными от абсолютных величин. Они выражают соотношения между абсолютными величинами и поэтому изменяются в зависимости от изменения абсолютных величин.

Однако количественное выражение относительных величин зависит не только от количественного различия между абсолютными величинами, которые сопоставляются с другими, но и от размера базы сравнения. Так, например, валовая продукция промышленности СССР в 1970 г. (374,3 млрд. руб.) составляет по сравнению с валовой продукцией 1969 г. (345,0 млрд. руб.) 108,5%, а по сравнению с валовой продукцией 1965 г. (229,4 млрд. руб.) и 1960 г. (157,4 млрд. руб.) — соответственно 163,2 и 237,8%. Из этого примера видно, что, чем меньше абсолютная величина, с которой производится сравнение, тем больше относительная величина, и наоборот, одна и та же абсолютная величина будет выражена различной относительной величиной в зависимости от размера базы сравнения. Поэтому одному и тому же проценту прироста может соответствовать различное абсолютное значение. Так, если процент прироста валовой продукции промышленности в 1970 г. по сравнению с 1960 г. составляет 1,574 млрд. руб.

$$\left(\frac{374,3 - 157,4}{157,4} \cdot 100 \right), \text{ по сравнению с } 1965 \text{ г. — } 2,294 \text{ млрд. руб.}$$

$$\left(\frac{374,3 - 229,4}{229,4} \cdot 100 \right), \text{ а по сравнению с } 1969 \text{ г. — } 3,45 \text{ млрд. руб.}$$

$$\left(\frac{374,3 - 345,0}{345,0} \cdot 100 \right).$$

Указанными особенностями взаимосвязи между абсолютными и относительными величинами и обуславливается необходимость

их комплексного использования в анализе. Взятые в отрыве одна от других, они не дают правильного представления об изучаемых явлениях и процессах.

Особенно серьезное внимание при исчислении относительных показателей следует уделять вопросу сопоставимости сравниваемых абсолютных величин. В первую очередь это относится к расчету относительных величин выполнения плана, динамики и сравнения. Сопоставимость абсолютных величин должна быть обеспечена в различных отношениях. Прежде всего необходима, чтобы все сравниваемые абсолютные величины характеризовали одно и то же явление и были однородны по содержанию и границам объекта, который они характеризуют. Так, сопоставляя производство электроэнергии за несколько лет или в нескольких странах, нельзя в одном случае включать, а в другом — не включать в общий итог электроэнергию, израсходованную на собственные нужды электростанций, электроэнергию, произведенную на передвижными электростанциями, электростанциями на судах, в поездах и т. п.

Несоблюдение этого условия может привести к ошибочным выводам при анализе общественных явлений и процессов. Одну из ошибок этого рода вскрыл в свое время В. И. Ленин. Народник Карышев, пытаясь доказать, что капитализм в России якобы не развивается, на основе взятых из разных источников сведений построил такой ряд динамики числа промышленных заведений в Европейской России: в 1879 г. — 27 986, 1890 г. — 21 124, 1894/95 гг. — 14 578. Отсюда Карышев делал вывод о сокращении числа промышленных предприятий в России и об отсутствии условий для развития капитализма.

В. И. Ленин, критикуя Карышева, писал, что «сочиненный г-ном Карышевым вывод получился оттого, что ученый профессор сравнивает совершенно несравнимые данные»¹. Несопоставимость данных была вызвана разноречивым толкованием понятий «фабрика» и «завод» в различные годы, в результате чего мелкие предприятия то учитывались, то не учитывались. Для получения сопоставимых данных В. И. Ленин выделил за каждый год предприятия с 16 и более рабочими и получил новый ряд динамики, начисто опровергающий выводы Карышева².

Таблица 5.5
ЧИСЛО ФАБРИК И ЗАВОДОВ
ЕВРОПЕЙСКОЙ РОССИИ

Годы	Всего	Имевших 16 и более рабочих
1879	27 986	4 551
1890	21 124	6 013
1894/95	14 578	6 659
		а без типографий 6 372

¹ Ленин В. И. Полн. собр. соч., т. 4, с. 12.

² См.: там же, т. 4, с. 13.

Таким образом, в действительности число фабрик и заводов в России увеличивалось, и притом довольно быстро.

Очень важным условием сопоставимости абсолютных величин является *одинаковая методология их исчисления*. Известно, что величина урожайности сельскохозяйственных культур может исчисляться либо в расчете на 1 га площади посева, которая имела в конце весеннего сева, либо в расчете на 1 га фактически убранной площади, т. е. за вычетом площади посевов, погибших летом. Поэтому при сопоставлении средней урожайности той или иной культуры за несколько лет или по отдельным районам необходимо во всех случаях сравнивать урожайность, исчисленную одинаковым методом, т. е. в расчете либо на 1 га посева, либо на 1 га фактически убранной площади.

Необходимо также следить за тем, чтобы все сравниваемые величины были выражены в одинаковых единицах измерения. Нельзя, например, продукцию одной ткацкой фабрики, выраженную в погонных метрах, сопоставлять с продукцией другой фабрики, выраженной в квадратных метрах, и т. п.

Сложнее решается вопрос о сопоставимости или несопоставимости абсолютных показателей в тех случаях, когда в течение периода, к которому они относятся, происходили административные территориальные изменения, изменения в подведомственности предприятий. В этих случаях вопрос о сопоставимости решается по-разному в зависимости от целей исследования. Допустим, что вследствие изменения административных границ или в связи с реорганизацией управления с 1 января 1972 г. в подчинение какому-либо министерству были переданы дополнительно 12 предприятий. Возникает вопрос: как обеспечить сопоставимость показателей выпуска продукции по данному министерству за годы до и после его реорганизации?

Ответ зависит от того, что именно нас в данном случае интересует. Если нас интересует министерство как хозяйственная система, то передача ему 12 предприятий не будет нарушать сопоставимости показателей. Если же нас интересуют успехи министерства в деле увеличения выпуска продукции и мы должны дать оценку его работы, то при сравнении показателей выпуска продукции их нужно рассчитывать за каждый год по сопоставимому числу предприятий, т. е. во всех случаях либо учитывать, либо не учитывать продукцию указанных 12 предприятий. Такое решение диктуется тем, что увеличение выпуска продукции за счет переданных предприятий нельзя считать достижением данного министерства.

Если же число подчиненных министерству предприятий возросло за счет ввода в действие новых предприятий, то это не нарушает сопоставимости данных, так как ввод в действие новых мощностей является одним из факторов роста производства.

В целях обеспечения сопоставимости в статистических справочниках и ежегодниках ряд показателей по нашей стране за 1913 г. приводится как в границах тогдашней России, так и в гра-

лициях СССР — довоенных и послевоенных. Например, в 1913 г. численность населения России в существовавших тогда границах составляла 165,7 млн. человек, а на территории, которую наша страна занимает в настоящее время, т. е. в современных границах СССР, численность населения была несколько меньше — 159,2 млн. человек. И если нас интересует увеличение численности населения СССР за счет естественного прироста (т. е. за счет того, что рождается людей больше, чем умирает), то при сравнении показателей численности населения мы должны устранить влияние других причин, в частности влияние изменения территории, и современную численность населения нашей страны сопоставлять с численностью того населения, которое проживало в 1913 г. на современной территории СССР.

Абсолютные показатели, выражающие размер явления за определенные промежутки времени, должны быть сопоставимы в отношении продолжительности периодов времени, к которым они относятся. Понятно, например, что данные о производстве продукции за три квартала несопоставимы с данными за год. Поэтому данные за неполный год (например, за 9 или 10 месяцев) сопоставляются с данными за соответствующий период прошлого года.

Многие явления имеют сезонный характер, т. е. их уровень из года в год в определенные месяцы повышается, а в другие — снижается¹. Это обстоятельство также может явиться причиной несопоставимости абсолютных величин. Известно, что цены колхозного рынка на молоко, фрукты и другие продукты летом ниже, чем зимой. Поэтому если сезонные колебания нас не интересуют, то для правильного представления об изменении цен нужно сопоставить уровни цен, приуроченные к одинаковой дате.

Строгое соблюдение требований сопоставимости особенно необходимо при сравнении показателей развития народного хозяйства СССР и капиталистических стран, так как методология исчисления одноименных показателей в советской и буржуазной статистике часто неодинакова.

Например, статистика США продукцию хлопкоочистительных предприятий относит к сельскому хозяйству, а не к промышленности, как это делается в СССР. Производство электроэнергии в США и других капиталистических странах учитывается по отпуску с шин, т. е. за вычетом той энергии, которая потреблена электростанциями на собственные нужды, тогда как в СССР учитывается вся выработка электроэнергии.

В наших статистических ежегодниках все показатели по СССР и США приводятся в сопоставимом виде — пересчитанными по единой методологии.

Вопрос о сопоставимости сравниваемых величин в каждом отдельном случае следует решать в зависимости от конкретных задач исследования.

¹ Подробнее о сезонности и ее изучении см. гл. IX.

4 ГРАФИЧЕСКИЕ СПОСОБЫ ИЗОБРАЖЕНИЯ АБСОЛЮТНЫХ И ОТНОСИТЕЛЬНЫХ ВЕЛИЧИН

В практике экономической работы как абсолютные, так и относительные статистические показатели часто изображаются графически. В предыдущей главе были рассмотрены общие вопросы графического способа изображения статистических данных и некоторые виды диаграмм. Выбор того или иного вида и конкретной формы графика определяется сущностью и познавательными задачами тех статистических показателей, которые должны быть изображены на графике. Рассмотрим некоторые специфические виды и формы графиков, применяемые при изображении абсолютных и различного вида относительных величин.

Графики выполнения плана

В практике экономической работы графики находят широкое применение для повседневного всестороннего контроля за ходом выполнения плана. В этих целях используются преимущественно различные ленточные (полосовые) и линейные графики.

Одним из видов линейных графиков выполнения плана является график нарастающих итогов. В прямоугольной системе координат на оси абсцисс откладываются отрезки, изображающие те календарные периоды (сутки, месяцы и т. п.), на которые делится весь плановый период. На вертикалях, проходящих через начало и конец планового периода, строятся сопряженные шкалы для нарастающих итогов (т. е. итогов с начала планового периода): одна шкала — для абсолютных величин, другая — для процентов выполнения плана за весь период. Сначала на график наносятся нарастающие итоги, предусмотренные планом. Если, например, график должен изображать выполнение годового плана, движущего помесячному развитию, тогда на вертикали, проходящей через конец января, откладывается январский план, на вертикали февраля — план двух первых месяцев, на ординате конца марта — план трех месяцев и т. д. Соединив отрезками нанесенные на вертикали точки (первая точка соединяется также с началом координат), получим плановую кривую нарастающих итогов — плановую кумуляту (в случае равномерной разработки годового плана по месяцам плановая кумулята будет представлять собой прямую линию). По мере выполнения плана на график таким же образом наносятся фактические нарастающие итоги с начала года (или другого планового периода). Для сравнения на графике можно построить такую же кумуляту за прошлый год.

Такой график дает наглядное представление о ходе выполнения плана за время, истекшее с начала планового периода, и о состоянии выполнения плана на данный момент. Так, график на рис. 5.1 показывает, что в январе план был недовыполнен; план февраля был выполнен (кривые идут параллельно), а план марта перевыполнен (фактическая кривая приближается к плановой). Однако январское недовыполнение в марте перекрыто еще не было

и только в апреле предприятие «вошло в график» и перевыполнило план первых четырех месяцев и т. д.

В тех случаях, когда необходимо контролировать выполнение плана одновременно по нескольким участкам производства и по двум-трем видам работ или продукции, большую помощь в оперативном контроле могут оказать ленточные диаграммы (см. рис. 5.2). По каждому участку и виду работ за 100% принимается плановый уровень на весь период. По мере выполнения плана каждая полоса систематически удлиняется, так что по состоянию на каждый данный момент времени видно, как обстоит дело на

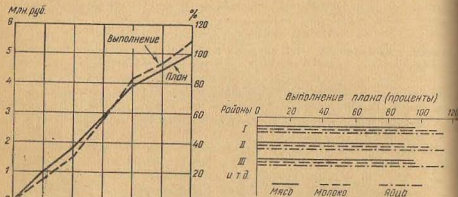


Рис. 5.1. График выполнения полу-годового плана (нарастающие итоги)

различных участках. График дает возможность видеть передовые и наиболее отстающие участки и своевременно принимать необходимые меры.

Недостаток такого рода графиков заключается в том, что они не показывают, как выполнялись плановые задания, установленные на более короткие периоды времени. Так, в нашем примере (рис. 5.2) из графика видно лишь выполнение плана, установленного на весь год, но неизвестно, как, например, был выполнен план продажи в III квартале, в том числе с учетом выполнения плана предыдущих кварталов. В связи с этим в практике применяются и более сложные графики, позволяющие дать ответ на такие вопросы.

Графики структуры Для графического изображения структуры явлений используются различные виды диаграмм: линейные, столбиковые, квадратные, круговые, в том числе секторные, и др. Выбор того или иного вида диаграммы зависит от того, для какой цели она строится, а также от характера изображаемых данных.

Чаще всего структура явления изображается в сравнении со структурной аналогичного явления в другом месте или в другое время, т. е. изображаются различия в структуре или структурные сдвиги. При таких сравнениях возможно построение структурных диаграмм двоякого рода. Если наряду с изображением структуры необходимо также изобразить графически и абсолютные размеры сравниваемых явлений, то строятся структурно-абсолютные диаграммы. Если же преследуется цель сопоставить только относительные величины структуры (удельные веса), то строятся структурно-относительные диаграммы.

Например, необходимо изобразить следующие данные.

Таблица 5.6

НАСЕЛЕНИЕ МИРА в 1919 и 1972 гг.

	1919 г.		1972 г.	
	Млн. человек	В процентах к итогу	Млн. человек	В процентах к итогу
Весь мир	1 777	100	3 750	100
в том числе:				
социалистические страны	138	7,8	1 223	32,6
остальные страны	1 639	92,2	2 527	67,4

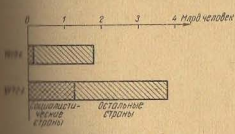


Рис. 5.3. Население мира в 1919 и 1972 гг. (полосовая диаграмма абсолютной структуры)

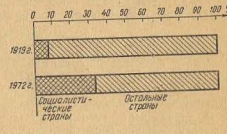


Рис. 5.4. Удельный вес населения социалистических стран в населении земного шара (процентом) полосовая диаграмма относительной структуры)

Изобразим эти данные в виде полосовой диаграммы абсолютной структуры (рис. 5.3). Диаграмма наглядно показывает рост численности населения земного шара, в том числе населения социалистических стран. Однако, чтобы оценить их долю в населении мира и проследить ее изменение, требуется некоторое напряжение воображения, так как при различной длине полос изменение доли непосредственно не воспринимается. Поэтому если основная цель диаграммы — показать изменение удельных весов, то целесообразнее построить диаграмму относительной структуры (рис. 5.4). На такой диаграмме легко воспринимаются удельные веса и их из-

менение, однако абсолютные данные могут быть показаны только числами.

Если наиболее существенным в диаграмме является сопоставление удельных весов, но в то же время нужно дать некоторое представление и об абсолютном увеличении целого и его составных частей, целесообразно использовать круговые диаграммы, разделенные на секторы (секторные диаграммы). Площади кругов и секторов будут в этом случае изображать абсолютные размеры целого и его частей, а центральные углы секторов — удельные веса. Для определения центральных углов могут быть использованы как абсолютные, так и относительные данные, пропорционально которым нужно разделить 360° . Так, в нашем примере угол сектора, изображающего население социалистического мира в 1972 г., равен: а) при расчете по абсолютным данным: $(360^\circ : 3750) \times 1223 = 118^\circ$; б) при расчете по процентам: $(360^\circ : 100) \times 32,6 = 117^\circ$.

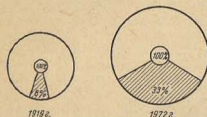


Рис. 5.5. Удельный вес населения социалистических стран в населении земного шара (процент; секторная диаграмма)

Площадь второго круга (1972 г.) должна быть больше площади первого круга (1919 г.) в 2,11 раза $(3750 : 1777)$, а радиус — в 1,45 раза (рис. 5.5).

Для графического изображения трех взаимосвязанных показателей, один из которых равен произведению двух других, русский статистик проф. В. Е. Варзар предложил использовать прямоугольную диаграмму, названную им «статистическим знаком». В настоящее время такие диаграммы часто называют знаками Варзара.

Простой знак Варзара строится в виде прямоугольника, основание которого пропорционально одному показателю-сомножителю, а высота — второму показателю-сомножителю. Тогда произведение этих показателей, т. е. третий показатель, будет изображаться площадью прямоугольника.

Например, нужно изобразить этим способом общее производство какой-либо продукции, ее производство на душу населения и среднюю численность населения. Взаимосвязь этих показателей можно представить так:

$$\left(\begin{array}{c} \text{Производство продукции} \\ \text{на душу населения} \end{array} \right) \times \left(\begin{array}{c} \text{Численность населения} \\ \text{(средняя)} \end{array} \right) = \left(\begin{array}{c} \text{Общее производ-} \\ \text{ство продукции} \end{array} \right).$$

Если отложить (в соответствии с принятыми масштабами) среднюю численность населения по основанию прямоугольника,

а производство продукции на душу населения — по его высоте, то площадь полученного прямоугольника будет изображать общее производство продукции.

При сравнении в пространстве или во времени прямоугольники размещаются на горизонтальной базовой линии либо рядом друг с другом (см. рис. 5.6), либо один на другом (чтобы совмещались вершины их левых нижних углов, которые служат началом отсчета показателей-сомножителей). Последний способ расположения особенно удобен при сравнении показателей за два периода времени. При размещении прямоугольников на горизонтальной линии по высоте следует откладывать тот показатель-сомножитель, сравнение величины которого является более важным.

Для изображения структуры (сложный знак Варзара) площадь прямоугольника может быть подразделена на части. Если по каждому сравниваемому объекту, кроме трех функционально связанных показателей, нужно изобразить еще несколько показателей, которые с основными непосредственно не связаны, под прямоугольниками помещаются дополнительные линейные или полосовые диаграммы (комбинированный знак Варзара).

Недостатком знака Варзара является то обстоятельство, что соотношение площадей прямоугольников оценить на глаз трудно, особенно если один из них имеет большее, чем у другого, основание, но меньшую высоту. Поэтому если наиболее важно сравнить показатели, представляющие произведение двух других, то лучше эти показатели изобразить в виде столбиковой диаграммы наряду со знаком Варзара.

Для графического изображения двух показателей, взаимосвязь которых является не функциональной, а корреляционной, используются особые диаграммы, которые рассматриваются в гл. VIII.

Распределение (размещение) явлений по территории изображается графически с помощью картодиаграмм и картограмм. Картодиаграмма — это контурная карта или план территории, на отдельные районы которой нанесены диаграммы в виде столбиков, кругов и т. п. На картодиаграмме обычно изображаются суммарные (итоговые) абсолютные величины: объем продукции, поголовье скота, жилищный фонд и т. п. Одновременно часто показывается и структура явления (в виде частей столбиков или секторов круга). На рис. 5.7 приведена картодиаграмма роста численности населения в союзных респуб-

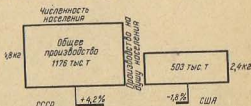


Рис. 5.6. Производство животного масла в СССР и США в 1972 г. (всего и в расчете на душу населения) и среднегодовые темпы прироста общего производства за 1951—1972 гг. (комбинированный знак Варзара)

ликах Закавказья с одновременным изображением распределения населения на городское и сельское.

На картограммах величина показателя может изображаться с помощью так называемых фоновых знаков — различной раскраски или штриховки территориальных единиц на контурной карте (фоновые картограммы), с помощью точек, которые наносятся на территорию соответствующих районов (точечные картограммы), а также путем проведения на контурной карте изолиний, т. е. линий, соединяющих места с одинаковой величиной показателя (изолинейные картограммы). В социально-экономической статистике чаще всего используются фоновые и точечные картограммы.

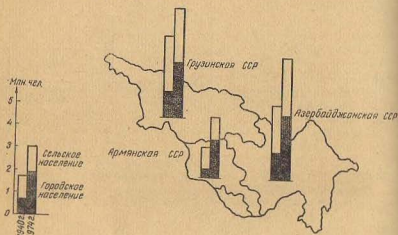


Рис. 5.7. Население союзных республик Закавказья в 1940 и 1974 гг. на 1 января (картограмма)

На фоновых картограммах обычно изображаются средние и относительные величины, например плотность населения, производство продукции на душу населения, урожайность и т. п.

При построении картограммы вначале выбираются территориальные единицы, характеризующиеся определенным показателем. Чаще всего при этом используются административно-территориальными единицами, к которым приурочены статистико-экономические показатели. Если общее число территориальных единиц на данной территории больше 6—8, то при построении фоновой картограммы их следует разгруппировать по величине изображаемого показателя на несколько групп, для каждой из которых устанавливается определенный тип раскраски или штриховки (число групп не должно превышать 6—8, так как при большем числе групп диаграмма потеряет наглядность из-за чрезмерной пестроты штриховки). Число групп и их границы устанавливаются в зависимости от конкретных особенностей материала. Для каждой группы стро-

ится шкала раскраски или штриховки. При этом необходимо соблюдать следующее правило: чем больше величина изображаемого показателя, тем интенсивнее должна быть для данной группы раскраска или штриховка.

Увеличение интенсивности окраски достигается путем перехода от светлых тонов какого-либо цвета к более темным тонам. Возможно также построение контрастной картограммы, на которой тонами одного цвета изображены значения показателя ниже среднего уровня, а тонами другого цвета — выше среднего уровня. В этом случае светлыми тонами каждого цвета изображаются величины, близкие к среднему уровню, а темными тонами — величины, сильно отклоняющиеся от среднего уровня.

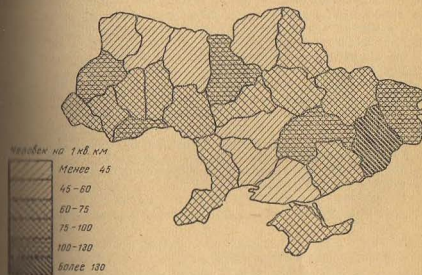


Рис. 5.8. Плотность населения областей УССР на 1 января 1973 г. (фоновая картограмма)

Повышение интенсивности штриховки может быть достигнуто различными способами: 1) переходом от пунктирных (точечных) линий к штрихам, а от штрихов к сплошным; 2) путем уменьшения расстояния между линиями; 3) путем утолщения линий; 4) переходом от штриховки в одном направлении к штриховке в двух и трех направлениях. Обычно эти способы комбинируют, например уменьшение расстояния между линиями сопровождается утолщением их.

Используются и более точные методы обеспечения нарастания интенсивности штриховки в соответствии с увеличением значения показателя. Один из них заключается в том, что при штриховке отдельных территориальных единиц в одном направлении расстояния между линиями берутся обратно пропорциональными величинам показателей. Однако этот способ применим лишь при значительной относительной колеблемости показателя. То же можно сказать и о втором способе (только квадратная штриховка сделана с таким расчетом, что на один квадратик на 1 см² было прямо пропорционально величинам изображаемых показателей).

В качестве иллюстрации построим фоновую картограмму плотности населения по областям УССР. Распределив области на

шесть групп по плотности населения на 1 января 1973 г., установленным шкалу штриховки (см. рис. 5.8).

Далее на контурной карте Украины, на которой нанесены границы областей, покроем территорию каждой области соответствующей штриховкой. В результате получим картограмму, изображенную на рис. 5.8. На картограмме отчетливо выделяются наиболее густонаселенные районы (на западе и на востоке республики) и районы с относительно низкой плотностью населения (на севере и юге).

На точечных картограммах величина показателя изображается соответствующим числом точек, каждая из которых принимается за определенное количество единиц измерения показателя. Так, чтобы изобразить численность населения по областям УССР, можно принять, например, что каждая точка соответствует 100 тыс. жителей. Если точки в пределах каждой территориальной единицы размещать равномерно, то число точек будет выражать (в соответствии с принятым масштабом) общую величину показателя, а их густота — среднюю плотность размещения явления по территории. Другой вариант построения заключается в том, что точки в пределах каждой территориальной единицы располагаются так, чтобы по возможности отразить фактическое местонахождение единиц объекта. Однако в этих случаях подсчитать число точек бывает трудно из-за значительного их сгущения на некоторых участках территории.

В практике используются и другие методы построения картограмм и картодиаграмм, в том числе различные их комбинации.

Глава VI

СРЕДНИЕ ВЕЛИЧИНЫ И ПОКАЗАТЕЛИ ВАРИАЦИИ

1. СУЩНОСТЬ И ЗНАЧЕНИЕ СРЕДНИХ ВЕЛИЧИН

Сущность и определение средней. Основные положения теории средних

Каждая однородная статистическая совокупность в области общественных явлений состоит из массы отдельных единиц, которые обладают индивидуальными особенностями и поэтому отличаются друг от друга по размеру количественных признаков. Так, отдельные токарки машиностроительного завода имеют неодинаковый тарифный разряд, получают различную заработную плату и т. д. Вместе с тем однородные статистические совокупности, именно вследствие своей однородности, обладают (в конкретных условиях места и времени) рядом типичных обобщающих характеристик, играющих исключительно важную роль в социально-экономическом анализе.

Например, нужно определить, на каком из двух заводов выше уровень заработной платы токарей. Для этого уровень заработной платы токарей по каждому заводу следует выразить одним числом и сравнить полученные величины. Ясно, что индивидуальная заработная плата отдельных токарей не может быть использована для такого сравнения, так как она различна в зависимости от ряда причин (квалификация, количество проработанного времени и др.). Чтобы сравнение было объективным, по каждому заводу необходимо учесть заработки всех токарей без исключения, т. е. нужно сравнивать *обобщающие показатели* уровня заработной платы. Однако непригодны для такого сопоставления и общие суммы заработной платы всех токарей завода, так как каждая такая сумма (фонд заработной платы) при прочих равных условиях тем больше, чем больше токарей на данном заводе. Таким образом, заработки отдельных токарей не годятся для сравнения потому, что не являются обобщающими показателями, а фонды заработной платы потому, что не выражают уровень заработной платы в расчете на одного токаря. Лишь разделив фонд заработной платы на число токарей и получив *среднюю заработную плату* по каждому заводу, можно сравнить полученные величины и определить, на каком предприятии *уровень заработной платы* токарей выше.

Следовательно, задача средней — одним числом охарактеризовать уровень признака у всех единиц однородной совокупности у которых размер признака варьирует, т. е. колеблется от одних единиц к другим. Свойство средней характеризовать не отдельные единицы, а выразить уровень признака в расчете на каждую единицу совокупности является ее отличительной особенностью. Эта особенность делает среднюю обобщающим показателем уровня варьирующего признака, т. е. показателем, который абстрагируется от индивидуальных различий размера признака у отдельных единиц однородной совокупности.

Таким образом, *под средней величиной в статистике понимают обобщающий показатель, который характеризует типичный уровень варьирующего признака в расчете на единицу однородной совокупности в конкретных условиях места и времени.*¹

Размеры варьирующего признака у отдельных единиц совокупности складываются под влиянием целого ряда условий и факторов. Одни из них могут быть общими, одинаковыми для всех единиц, а другие — различными. Эти различия и порождают вариацию признака. Так, например, для всех токарей завода, имеющих данный тарифный разряд, установлена единая тарифная ставка, а при выполнении одной и той же работы — одинаковые нормы выработки. Однако размеры месячной заработной платы у отдельных токарей данного разряда, даже выполнявших одинаковую работу, как правило, неодинаковы, так как наряду с общими условиями на величину заработка влияют также индивидуальные и чисто случайные факторы: состояние станка, интенсивность труда, количество отработанного за месяц рабочего времени и др. *Средняя отражает то общее, что скрывается в каждой единице однородной совокупности.* Она улавливает общие черты, общую закономерность массовых общественных явлений, которые проявляются в силу закона больших, или, как называет его К. Маркс, закона средних чисел. Вместе с тем в средней величине погашаются индивидуальные особенности отдельных единиц совокупности. Так, установив, что труд, овеществленный в стоимости товара, есть труд среднего общественного качества, К. Маркс указывал, что хотя в каждой отрасли промышленности индивидуальные рабочие отклоняются в той или иной степени от среднего рабочего, но эти отклонения, или, как их называют математики, «погрешности», взаимно погашаются и уничтожаются, если взять значительное число рабочих.

Основное условие научного использования средних в статистике заключается в том, чтобы средняя характеризовала такую совокупность единиц, которая в существенном отношении, в первую очередь в отношении осредняемого признака, является *качественно однородной*. Только при этом условии средняя действительно будет отражать то общее, что существенно и характерно для данной совокупности, и выявить типичный размер признака. Если же совокупность качественно неоднородна, например состоит из единиц, относящихся к различным социально-экономическим

типам, или из групп, существенно отличающихся друг от друга в отношении осредняемого признака, то средняя, исчисленная для такой разнородной совокупности, будет искажать действительное положение вещей. Такого рода средние имеют чисто фиктивный характер, они не только не обогащают, но, наоборот, затрудняют познание реальной действительности.

Критикуя разработку материалов кустарной переписи 1894—1900 гг. в Пермской губернии русскими земскими статистиками, В. И. Ленин отмечал, что, соединяя крупные и мелкие заведения вместе, они выводят фиктивные «средние» цифры, которые не только не дают понятия о действительности, но, наоборот, затуманивают кардинальные различия и изображают однородным нечто совершенно разнородное. Эти «средние», как кривое зеркало, только искажали содержание процесса развития промышленности в России.

К таким же фиктивным «средним» земские статистики прибегали и при характеристике развития русской деревни. Вместо того чтобы четко ограничить отдельные экономические типы хозяйств и по отношению к каждой такой однородной группе вычислять средние показатели, они выводили средние величины по всей площади, скота и т. д. на одно домохозяйство в целом по уезду или губернии. В. И. Ленин по этому поводу писал, что обработка статистических данных земскими статистиками сводится к одному сплошному и невероятному злоупотреблению средними величинами и с тонкой иронией отмечал, что «стоит только пользоваться всегда и исключительно «средними» данными о крестьянском хозяйстве, — и все «превратные идеи» о разложении крестьянства окажутся раз навсегда изгнанными»¹.

Для выделения качественно однородных совокупностей в статистике используется метод группировок². Это определяет тесную органическую связь метода средних величин с методом группировки: только после того как в результате качественного социально-экономического анализа образованы типические однородные группы (совокупности), имеет смысл характеризовать их средними величинами изучаемых признаков. Вне метода группировки, по отношению к массе неоднородных единиц, средние величины приобретают огульный, фиктивный характер, теряют социально-экономическую значимость.

В ходе дальнейшего анализа качественно однородная совокупность может быть вновь подразделена на характерные в том или ином отношении группы. В этом случае общая средняя, исчисленная для такой совокупности, должна быть дополнена групповыми средними, характеризующими отдельные группы. Также представляется целесообразным дополнять средние индивидуальные показатели. Например, общую по колхозам области среднюю урожайность овощных культур дополнять показателями

¹ Ленин В. И. Полн. собр. соч., т. 3, с. 162—163.
² Там же, IV, с. 78 и далее.

средней урожайности на поливных и неполивных землях, а также показателями средней урожайности овощных культур, достигнутыми передовыми бригадами отдельных колхозов.

Необходимость в ходе анализа дополнения и корректировки общей средней средними по группам, подгруппам и индивидуальным показателям не означает, конечно, что общую среднюю можно отождествлять с огульной, фиктивной средней. Чтобы общая средняя не была фиктивной, ее нужно исчислять для такой совокупности, которая действительно, а не только формально, имеет те или иные общие существенные черты и свойства. Существенность этих свойств устанавливается путем конкретного социально-экономического анализа.

Замечательные примеры научного применения метода средних в сочетании с методом группировок, исчисления частных и групповых средних, относящихся к качественно однородной массе, дают работы В. И. Ленина, материалы партийных съездов и Пленумов ЦК КПСС.

Использование средних в экономическом анализе

В условиях социалистического планового хозяйства средним величинам принадлежит важное место в экономических исследованиях. Они используются для характеристики важнейших закономерностей развития общественных явлений, существующих и намечающихся тенденций и т. п. Например, закономерность неуклонного роста производительности труда, свойственная социалистической промышленности, находит выражение в статистических показателях роста средней производительности труда. Имеющиеся в планах и отчетности промышленных предприятий показатели производительности труда, затрат на рубль товарной продукции, себестоимости единицы продукции и ряд других по самой своей природе являются средними показателями.

В экономике сельского хозяйства закономерность повышения урожайности, свойственная социалистическому сельскому хозяйству, находит выражение в статистических показателях роста средней урожайности. Следует заметить, что урожайность по участку, бригаде, колхозу, а тем более по району, области есть не что иное, как средняя урожайность, так как величина эта обобщающая, отнесенная к площади, на которой можно встретить различную урожайность на каждом квадратном метре. Точно так же дело обстоит и с показателями продуктивности животноводства. Надой молока на одну корову, настриг шерсти в расчете на одну овцу и т. д. — все это обобщающие, средние по своей сущности показатели.

В экономическом анализе средним величинам принадлежит важная роль во вскрытии неиспользуемых внутрипроизводственных резервов, в обосновании экономической эффективности внедрения новых материалов, новой техники и технологии. Сопоставляя, например, среднюю себестоимость или среднюю трудоемкость (затраты труда) единицы продукции в условиях применения раз-

ных материалов или различной технологии, можно судить об их сравнительной экономической эффективности. Известно, что средняя трудоемкость 1 м² жилой площади в крупноблочных и крупнопанельных домах составляет примерно 2,5—3,5 человеко-дня, а в кирпичных домах — 4,5—6 человеко-дней. Следовательно, строительство домов из крупных панелей и блоков дает большую экономическую затрат труда.

Наглядное представление о резервах роста производства дает сопоставление показателей, достигнутых на передовых участках, со средними показателями по области, республике или стране. Например, горняки шахты «Наторная» (Новокузнецк) развернули социалистическое соревнование за добычу 100 т угля в месяц на каждого шахтера в 1973 г., решающем году девятой пятилетки, тогда как в среднем по стране этот показатель намного ниже.

Однако как бы эффектно ни выглядели достижения передового производства на фоне средних по стране показателей, нельзя забывать о том, что за средними скрываются также низкие показатели. Поэтому средними характеристиками нужно пользоваться с большой осторожностью, не преувеличивая их значения, так как средняя, являясь обобщающим показателем, погашает те количественные различия варьирующего признака, которые проявляются у отдельных единиц совокупности. Общие средние, за которыми не видно лица отдельных предприятий, колхозов, строков и т. д., нередко порождают самоуспокоенность и благодушие у хозяйственных руководителей, что приносит большой вред. Вот почему советская статистика исходит из того, что средние должны опираться на конкретными индивидуальными данными. С помощью индивидуальных данных устанавливается и обобщается опыт передовиков, но с помощью этих же данных выявляются и отстающие участки, плохая работа.

2. СПОСОБЫ ВЫЧИСЛЕНИЯ СРЕДНИХ

Исходные основания расчета средних

Вычисление средних в статистике существенно отличается от их вычисления в математике. В математике рассматриваются различные формы средней (арифметическая, геометрическая, гармоническая и др.), их свойства и способы расчета. При этом выделяются соотношения различных средних, полученных из одного и того же ряда чисел (вариантов признака).

Например, имеется следующий ряд чисел: 1, 2, 3. Средняя арифметическая этих чисел будет:

$$\bar{x}_a = \frac{1+2+3}{3} = \frac{6}{3} = 2;$$

средняя геометрическая:

$$\bar{x}_{гк} = \sqrt[3]{1 \times 2 \times 3} = 1,817$$

и средняя гармоническая:

$$\bar{x}_{\text{гр}} = \frac{3}{1 + \frac{1}{2} + \frac{1}{3}} = \frac{3}{\frac{11}{6}} = \frac{18}{11} = 1,636.$$

Отсюда соотношение исчисленных средних выразится следующей цепью неравенств:

$$\bar{x}_a > \bar{x}_{\text{гн}} > \bar{x}_{\text{гр}}.$$

По-иному обстоит дело в статистике. Изучая количественную сторону массовых общественных явлений в неразрывной связи с их качественной стороной, статистика должна находить такие признаки, которые являлись бы типичными характеристиками признаков изучаемых явлений. В соответствии с задачей, стоящей в каждом конкретном случае, для этого должно быть найдено одно-единственное значение средней путем использования определенной ее формы. Так, в результате изучения уровня заработной платы рабочих предприятия должна быть получена его обобщающая характеристика в виде средней заработной платы. При этом для имеющихся исходных данных эта средняя должна иметь вполне определенное значение, быть однозначной. Поэтому было бы неверно для решения одной задачи к одним и тем же исходным данным применять различные формы средних и получать соответственно различные числовые значения средней.

Все это определяет особую важность правильного выбора формы средней. Форма средней в статистике подчинена социально-экономическому содержанию изучаемых явлений и обусловлена существующими между ними объективными взаимосвязями. Выражением взаимосвязей явлений и их признаков являются взаимосвязи характеризующих их статистических показателей.

Поскольку средняя характеризует уровень признака в расчете на единицу совокупности, постольку взаимосвязь между средней и показателями, от которых она непосредственно зависит, как правило, может быть выражена в виде кратного отношения двух итоговых показателей суммарного, объемного характера. В простейших, наиболее часто встречающихся случаях числитель такого отношения представляет собой общую сумму значений осредняемого признака у всех единиц совокупностей, т. е. общий объем признака, а знаменатель — общее число единиц совокупности, в расчете на каждую из которых вычисляется средняя¹. Например, средний размер семьи (по состоянию на какую-либо дату) выражается как отношение суммы размеров всех отдельных семей, т. е. общего количества их членов, к числу этих семей:

$$\text{Средний размер семьи} = \frac{\text{Общее количество членов всех семей}}{\text{Число семей}}.$$

¹ Более сложные случаи рассматриваются в гл. IX.

Приведем еще несколько аналогичных примеров:

$$\text{Время затраты времени на единицу продукции} = \frac{\text{Общие затраты времени на всю продукцию}}{\text{Количество единиц продукции}}.$$

$$\text{Средняя урожайность с одного гектара} = \frac{\text{Общий урожай (валовой сбор)}}{\text{Количество гектаров посева}}.$$

$$\text{Средний размер вклада в сберегательные кассы (взвешиваний остаток вкладов на ту или иную дату)} = \frac{\text{Общая сумма всех вкладов}}{\text{Число вкладов}}.$$

Подобного рода соотношения, выражающие смысл средних величин и их зависимость от других показателей, являются исходной базой расчета и критерием правильности выбора формы средней.

Возможны и такие случаи, когда осредняемый показатель сам по себе, независимо от процесса осреднения его значений, уже представляет собой относительную величину, т. е. отношение двух других показателей, например:

$$\text{Процент выполнения плана} = \frac{\text{Фактический уровень}}{\text{Плановый уровень}} \times 100.$$

В этих случаях исходная база расчета средней выражается в виде отношения тех же показателей, каждый из которых охватывает всю совокупность единиц, для которых исчисляется средняя.

$$\text{Средний процент выполнения плана выпуска продукции по нескольким заводам} = \frac{\text{Общий фактический объем продукции всех заводов}}{\text{Общий объем продукции этих заводов по плану}} \times 100.$$

Таким образом, средняя величина относительного показателя представляет собой этот относительный показатель, исчисленный в целом по всей совокупности.

Итак, первый этап исчисления средней в статистике состоит в нахождении исходной базы ее расчета в виде отношения связанных с ней суммарных итоговых показателей. После этого задача сводится к определению числовых значений этих показателей на основе имеющихся данных. Порядок дальнейших расчетов и форма средней зависят целиком от исходного отношения суммарных показателей (т. е. от исходной базы расчета средней), с одной стороны, и от того, какими данными для расчета средней мы располагаем, — с другой.

В простейшем случае — при наличии обоих суммарных итоговых показателей, отношением которых выражается средняя, — расчет ее производится непосредственно путем деления этих итоговых показателей. Так обычно исчисляется средняя по данным статистической отчетности. Например, по данным отчета промышленного предприятия о выполнении плана по труду средняя заработная плата, приходящаяся на одного рабочего, вычисляется

путем деления фонда заработной платы рабочих на их среднесписочную численность.

Во многих случаях, однако, один из суммарных итоговых показателей отсутствует, так что расчет средней приходится производить на основе первичных данных о значении осредняемого признака у каждой отдельной единицы или группы единиц. В этих случаях средняя величина должна исчисляться так, чтобы при замене каждого варианта осредняемого показателя средней величиной оставался без изменения некоторый фактический итоговый показатель, связанный с осредняемым. Этот показатель называется определяющим, так как его взаимосвязь с осредняемым показателем определяет форму средней.

Средняя арифметическая

Наиболее распространенной формой средней, применяемой в экономических расчетах, является средняя арифметическая. Вывод формулы и расчет средней арифметической покажем на следующем примере. Допустим, 10 фрезеровщиков изготовили за смену следующее количество шестерен:

Порядковый номер фрезеровщика	I	II	III	IV	V	VI	VII	VIII	IX	X
Изготовлено шестерен — штук	43	46	46	38	44	49	43	51	44	46

Нужно найти среднюю выработку на одного рабочего за смену. Исходная база ее расчета имеет такой вид:

$$\text{Средняя выработка на одного рабочего} = \frac{\text{Общая выработка всех рабочих}}{\text{Число рабочих}}$$

Общая выработка всех рабочих равна сумме сменных выработок всех 10 рабочих. Следовательно,

$$\begin{aligned} \text{Средняя выработка на одного рабочего} &= \frac{43 + 46 + 46 + 38 + 44 + 49 + 43 + 51 + 44 + 46}{10} = \\ &= \frac{450}{10} = 45 \text{ (штук)}. \end{aligned}$$

Характеризуя типичный размер варьирующего признака, средняя неизбежно совпадает с одним из вариантов. Так, в нашем примере средняя выработка составляет 45 шестерен, хотя ни один из рабочих такого количества шестерен не изготовил.

Обозначим выработку отдельных рабочих через x_1, x_2, \dots, x_n , а среднюю выработку — через \bar{x} . Определяющим показателем является в данном случае общая выработка всех рабочих, равная $x_1 + x_2 + \dots + x_n$. Величина этого показателя должна остаться без изменения при замене всех вариантов их средним значением. Следовательно, $x_1 + x_2 + \dots + x_n = \bar{x} + \bar{x} + \dots + \bar{x}$. Это уравнение

средней. Заменяв в нем сложенные одинаковых слагаемых умножением, получим

$$x_1 + x_2 + \dots + x_n = n\bar{x}, \quad \sum x = n\bar{x}$$

откуда

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n},$$

или в более простой записи:

$$\bar{x} = \frac{\sum x}{n}, \quad (6.1)$$

где \sum — знак суммирования.

Средняя, исчисленная по этой формуле, называется простой средней арифметической, так как она получается путем простого суммирования всех вариантов осредняемого показателя и деления полученной суммы на общее число всех вариантов.

Следовательно, если имеются первичные абсолютные данные о величине признака у каждой отдельной единицы и нужно исчислить среднюю величину признака в расчете на эту же единицу, то применяется формула простой средней арифметической.

В нашем примере некоторые варианты выработки (43, 44 и 46) повторяются несколько раз. Взяв только различные по величине варианты и указав число случаев повторения каждого из них (частоты), получим следующий вариационный ряд (гр. 1 и 2 табл. 6.1).

Таблица 6.1

РАСПРЕДЕЛЕНИЕ РАБОЧИХ ПО ЧИСЛУ ИЗГОТОВЛЕННЫХ ШЕСТЕРЕН (к расчету средней арифметической взвешенной)

Исходные данные		Расчет
выработка шестерен за смену одним рабочим (варианты) x	число рабочих, имеющих данную выработку (частоты) m	общая выработка шестерен xm
1	2	3
38	1	38
43	2	86
44	2	88
46	3	138
49	1	49
51	1	51
Итого	10	450

Чтобы по этим данным определить среднюю выработку на одного рабочего, нужно, как и ранее, общую выработку разделить на общую численность рабочих. Однако порядок расчета будет

инии. Для получения общей выработки всех 10 рабочих уже недостаточно просто сложить различные варианты выработки, нужно также учесть, сколько раз встречается каждый вариант, т. е. сколько рабочих имеют данную величину сменной выработки. Следовательно, чтобы подсчитать общую выработку всех рабочих, нужно каждый вариант выработки умножить на число рабочих, имеющих данную выработку, а затем сложить полученные результаты (см. гр. 3). Разделив полученную таким путем общую выработку на число рабочих, найдем среднюю выработку на одного рабочего:

$$\bar{x} = \frac{450}{10} = 45 \text{ (штук)}.$$

Обозначая различные варианты выработки через x_1, x_2, \dots, x_n , а числа рабочих, имеющих соответствующую выработку, через m_1, m_2, \dots, m_n , приходим к следующей формуле:

$$\bar{x} = \frac{x_1 m_1 + x_2 m_2 + \dots + x_n m_n}{m_1 + m_2 + \dots + m_n} = \frac{\sum x m}{\sum m}. \quad (6.2)$$

Средняя из вариантов, которые повторяются различное число раз, или, как говорят, имеют различный вес, называется взвешенной. В нашем примере мы получили формулу взвешенной средней арифметической.

Как видно из формулы (6.2), средняя арифметическая взвешенная равна сумме попарных произведений вариантов (x) на соответствующие им веса (m), деленной на сумму весов.

Данную формулу в нашем примере можно получить также из уравнения средней. Величина определяющего показателя (общей выработки) не должна измениться при замене вариантов их средней величиной. Следовательно,

$$x_1 m_1 + x_2 m_2 + \dots + x_n m_n = \bar{x} m_1 + \bar{x} m_2 + \dots + \bar{x} m_n = \bar{x} (m_1 + m_2 + \dots + m_n).$$

Отсюда

$$\sum x m = \bar{x} \sum m \text{ и } \bar{x} = \frac{\sum x m}{\sum m}.$$

Из рассмотренного примера можно сделать следующий вывод. В качестве весов (m) средней арифметической взвешенной используютсялагаемые показатели, находящегося в знаменателе исходного отношения, т. е. исходной базы расчета средней. Так, в рассмотренном примере весами являлись числа рабочих, имеющих одинаковую выработку, а их сумма (общая численность рабочих) представляла собой показатель, находящийся в знаменателе исходной базы расчета средней выработки.

В данном случае в качестве весов выступали частоты вариационного ряда, так как именно они являлись слагаемыми знаменателя исходной базы расчета средней. Не следует, однако, отождествлять веса взвешенной средней арифметической с частотами вариационного ряда. Частоты используются в качестве весов лишь при том условии, что их сумма, т. е. численность совокуп-

ности, подвергнутой группировке при построении вариационного ряда, представляет собой знаменатель исходного отношения. Между тем это имеет место далеко не всегда. Пусть, например, нужно определить в целом по району среднюю урожайность озимой пшеницы по следующим данным.

Таблица 6.2

РАСПРЕДЕЛЕНИЕ КОЛХОЗОВ ПО ВЕЛИЧИНЕ УРОЖАЙНОСТИ ПШЕНИЦЫ

(в расчете средней арифметической взвешенной, когда частоты не являются весами)

Исходные данные	Расчет		
	число колхозов, получающих такую среднюю урожайность (частоты)	посевная площадь пшеницы в этих колхозах (га) (веса— m)	валовой сбор (ц) xm
урожайность по колхозу (варианты— x)			
18	1	250	4 500
20	3	1 500	30 000
22	1	750	16 500
Итого по району	5	2 500	51 000

Исходная база расчета средней урожайности выражается так:

$$\text{Средняя урожайность с 1 гектара} = \frac{\text{Валовой сбор (общий урожай)}}{\text{Количество гектаров посевной площади}}$$

Чтобы найти в целом по району валовой сбор пшеницы, нужно по каждой группе колхозов умножить полученную с 1 га среднюю урожайность (x) на размер посевной площади в гектарах (m), а затем сложить полученные результаты. Разделив далее найденную сумму ($\sum xm$) на общий размер посевной площади ($\sum m$), получим искомую среднюю урожайность с 1 га в целом по району:

$$\bar{x} = \frac{\sum x m}{\sum m} = \frac{18 \times 250 + 20 \times 1 500 + 22 \times 750}{250 + 1 500 + 750} = \frac{51 000}{2 500} = 20,4 \text{ (ц с 1 га)}.$$

В данном случае, как и в предыдущем примере, в качестве весов средней арифметической выступают слагаемые показатели, находящегося в знаменателе исходного отношения. Но этими слагаемыми являются не частоты вариационного ряда, т. е. не числа колхозов, получивших данную урожайность, а посевные площади этих колхозов. Это объясняется тем, что при построении вариационного ряда группировались колхозы, средняя же урожайность исчисляется в расчете на один гектар, а не на один колхоз.

Обобщая сказанное, приходим к такому выводу: при расчете средних величин по данным вариационных рядов можно пользоваться формулой средней арифметической, принимая в качестве весов показатели, находящиеся в знаменателе исходного отношения (базы расчета средней).

Особенность последнего примера состоит в том, что варианты урожайности, на основе которых исчислялась средняя урожайность в целом по району, сами уже являлись средними величинами. Так, в колхозе, где было посеяно 250 га пшеницы, урожайность, конечно, не была одинаковой на каждом из этих 250 га и лишь в среднем она составила 18 ц с 1 га и т. д. По отношению к средней урожайности по району в целом, которая является общей средней, урожайность по каждой группе колхозов представляет собой групповую, или частную, среднюю.

С необходимостью исчисления общей средней (\bar{x}) на основе имеющихся частных (групповых) средних (обозначим их через x_i) в экономической работе приходится сталкиваться очень часто. В нашем примере с урожайностью такая необходимость возникла применительно к вариационному ряду. Аналогичный случай может иметь место и для атрибутивного ряда распределения, если каждая его группа характеризуется средней или относительной величиной какого-либо признака. Допустим, что на основе следующих данных (гр. 1—3 табл. 6.3) нужно определить средний процент выполнения плана в целом по всем бригадам.

Таблица 6.3

К РАСЧЕТУ ОБЩЕЙ СРЕДНЕЙ ИЗ ГРУППОВЫХ СРЕДНИХ

	Исходные данные			Расчет
	число бригад	план выпуска продукции (тыс. руб.) m_i	средний процент выполнения плана x_i	фактический выпуск продукции (тыс. руб.) $\frac{x_i m_i}{100}$
А	1	2	3	4
Бригады коммунистического труда	2	150	116	174
Бригады, борющиеся за присвоение звания бригад коммунистического труда	6	350	108	378
Итого	8	500	$\frac{552}{500} \times 100 = 110,4$	552

При расчете нужно исходить из того, что относительная величина выполнения плана есть отношение фактического уровня к плановому:

$$\text{Процент выполнения плана} = \frac{\text{Фактический уровень}}{\text{Плановый уровень}} \times 100.$$

Чтобы в соответствии с этой исходной базой найти фактический выпуск продукции по каждой группе бригад, нужно план выпуска умножить на процент выполнения плана и разделить на 100 (это сделано в гр. 4).

Фактический выпуск продукции всеми бригадами в целом (числитель исходного отношения) составляет, следовательно, 552 тыс. руб. $\left(\sum \frac{x_i m_i}{100}\right)$. Разделив эту величину на общую сумму выпуска продукции по плану (знаменатель исходного отношения $\sum m_i$) и умножив на 100, найдем процент выполнения плана всеми бригадами в целом, т. е. средний процент выполнения плана.

$$\bar{x} = \frac{\sum \frac{x_i m_i}{100}}{\sum m_i} \cdot 100 = \frac{552}{500} \times 100 = 110,4\%.$$

Нетрудно видеть, что использованная при этом формула (после сокращения на 100) является обычной формулой взвешенной средней арифметической, в которой роль вариантов осредняемого признака выполняют частные средние:

$$\bar{x} = \frac{\sum \frac{x_i m_i}{100}}{\sum m_i} \cdot 100 = \frac{\sum x_i m_i}{\sum m_i}. \quad (6.3)$$

Следовательно, общую среднюю можно исчислить как среднюю арифметическую из частных средних, весами которых являются соответствующие части знаменателя исходного отношения.

Подведем итоги. В рассмотренных примерах для расчета средней мы располагали вариантами осредняемого признака (x) и соответствующими слагаемыми (m) того показателя, который находится в знаменателе исходного отношения. Эти слагаемые знаменателя использовались в качестве весов. Величина же показателя, находящегося в числителе исходного отношения, равная сумме произведений вариантов на веса ($\sum x m$), непосредственно не была известна. Причем эта величина не должна была меняться при замене вариантов их средней величиной. В результате во всех этих случаях исчислялась взвешенная средняя арифметическая. Это позволяет сделать вывод об условиях и границах ее применения.

Взвешенная средняя арифметическая применяется в тех случаях, когда показатель, находящийся в числителе исходного отношения средней, непосредственно неизвестен, а его величина равна сумме произведений вариантов осредняемого признака (x) на соответствующие слагаемые знаменателя исходного отношения (веса — m).

Взвешенная средняя арифметическая используется и в тех случаях, когда варианты выражены не в дискретной форме, а в виде интервалов (от — до), т. е. для интервальных вариационных рядов. Например, нужно определить средний стаж ра-

боты по следующим данным:

Таблица 6.4
РАСПРЕДЕЛЕНИЕ РАБОЧИХ ПО СТАЖУ РАБОТЫ

Группы рабочих по числу лет работы на заводе	Число рабочих
До 1	16
1—3	48
3—5	59
5—10	50
10—20	19
20 и более	8
Итого	200

Исходной базой для расчета среднего стажа работы является отношение общего числа проработанных всеми рабочими человеко-лет к общей численности рабочих:

$$\text{Средний стаж работы (число лет работы в среднем на одного рабочего)} = \frac{\text{Общее число проработанных всеми рабочими человеко-лет}}{\text{Число рабочих}}$$

Чтобы найти число проработанных рабочими человеко-лет, нужно по каждой группе рабочих умножить стаж работы на число рабочих; имеющих этот стаж, и сложить полученные результаты. Обозначив варианты стажа работы через x , а число рабочих через m , в соответствии с исходным отношением получим формулу средней арифметической взвешенной:

$$\bar{x} = \frac{\sum xm}{\sum m}$$

Эта формула предполагает наличие вариантов x в дискретной форме, тогда как в интервальном ряду для каждой группы указаны лишь нижняя и верхняя границы интервала. В связи с этим возникает вопрос о том, какую же величину варианта x нужно применять в формуле в качестве множителя по каждой группе, т. е. для каждого интервала.

Исходя из предположения, что значения признака у отдельных единиц каждой группы распределяются внутри интервала более или менее равномерно, в качестве множителя для каждого интервала обычно берется его центр, т. е. середина между нижней и верхней границами интервала. Она находится как простая средняя арифметическая из этих границ. Так, в нашем примере середина второго интервала $(1+3):2=2$ годам, середина третьего интервала $(3+5):2=4$ годам и т. д.

Первый и последний интервалы в нашем примере являются открытыми: у первого интервала указана только верхняя его граница, а у последнего — только нижняя. Для первого интервала нижней границей является в данном случае нуль, так как стаж работы может быть как угодно мал. Отсюда середина первого

интервала $(0+1):2=0,5$ года. Верхней границей последнего интервала можно было бы считать 40 лет, так как больший стаж работы встречается очень редко. Однако стаж более 30 лет также встречается относительно редко. Поэтому верхнюю границу последнего интервала можно установить с таким расчетом, чтобы величина этого интервала оказалась равной величине соседнего с ним интервала. Такой принцип установления неизвестных границ открытых интервалов применяется обычно в тех случаях, когда невозможно установить эти границы исходя из характера показателя и конкретных условий. В соответствии с этим принципом величина последнего интервала должна быть равна величине предпоследнего интервала, т. е. $20-10=10$ годам. Следовательно, верхняя граница последнего интервала $20+10=30$ годам, а середина — $(20+30):2=25$ годам.

Дальнейшие расчеты по определению среднего стажа работы могут быть сведены в следующую таблицу.

Таблица 6.5

К РАСЧЕТУ СРЕДНЕЙ АРИФМЕТИЧЕСКОЙ ПО ДАННЫМ ИНТЕРВАЛЬНОГО ВАРИАЦИОННОГО РЯДА

Исходные данные		Расчет	
стаж работы (лет)	число рабочих m	середина интервала x	проработано (человеко-лет) xm
До 1	16	0,5	8
1—3	48	2	96
3—5	59	4	236
5—10	50	7,5	375
10—20	19	15	285
20 и более	8	25	200
Итого	200	—	1 200

Подставляя итоговые данные таблицы в формулу взвешенной средней арифметической, найдем средний стаж работы:

$$\bar{x} = \frac{\sum xm}{\sum m} = \frac{1 200}{200} = 6 \text{ (лет)}.$$

Необходимо иметь в виду, что расчет средней по данным интервального вариационного ряда является приближенным. Это объясняется тем, что в качестве множителя x для каждого интервала используется его середина, которая может отличаться от действительной средней величины признака в данной группе, если варианты в пределах интервала распределяются неравномерно. При наличии открытых интервалов к этому добавляются неточности, связанные с условным установлением неизвестных границ. Поэтому рассмотренный метод расчета средней для интервального ряда следует применять лишь в тех случаях, когда отсутствуют как данные о значении признака у каждой отдельной

единицы, так и данные об общем объеме признака для всей совокупности в целом. При наличии же таких данных точное значение средней величины может быть получено гораздо проще. Например, наряду с интервальным рядом распределения рабочих по размеру выработки продукции имеются также данные об общей выработке продукции всеми рабочими. В этом случае для вычисления средней выработки на одного рабочего нет необходимости использовать данные вариационного ряда, так как, разделив общую выработку на число рабочих, получим величину средней выработки без всякой погрешности.

Основные свойства средней арифметической. Упрощенные способы ее вычисления

Средняя арифметическая обладает рядом свойств, которые могут быть использованы для упрощения ее вычисления и в других целях. Рассмотрим основные свойства средней арифметической.

1. Алгебраическая сумма отклонений вариантов от средней арифметической равна нулю:

$$\sum (x - \bar{x}) m = 0. \quad (6.4)$$

Пусть дан вариационный ряд, в котором варианты принимают значения x_1, x_2, \dots, x_n и веса соответственно m_1, m_2, \dots, m_n . Средняя арифметическая для такого ряда по формуле (6.2) равна $\bar{x} = \frac{\sum x m}{\sum m}$, откуда следует, что

$$\sum x m = \bar{x} \sum m. \quad (6.5)$$

Обозначим разность между вариантами и средней δ (греческая буква дельта малая). Тогда алгебраическая сумма отклонений вариантов от средней арифметической составит:

$$\sum \delta m = \sum (x - \bar{x}) m = \sum x m - \bar{x} \sum m,$$

что согласно формуле (6.5) равно нулю.

Это свойство имеет значение при расчете показателей вариации (см. ниже).

2. Величина средней арифметической не изменится, если все каждого варианта умножить (или разделить) на одно и то же число.

Возьмем вариационный ряд, для которого по формуле (6.2) средняя арифметическая равна $\bar{x} = \frac{\sum x m}{\sum m}$. Умножим теперь все веса этого ряда на число a . Тогда средняя арифметическая будет

$$\bar{x} = \frac{\sum x a m}{\sum a m} = \frac{a \sum x m}{a \sum m} = \frac{\sum x m}{\sum m} = \bar{x}.$$

Следовательно, значение средней осталось без изменения. Из этого свойства вытекают определенные следствия.

А. Если веса всех вариантов равны между собой, то взвешенная средняя равна простой средней:

$$\bar{x} = \frac{x_1 a + \dots + x_n a}{a + \dots + a} = \frac{a (x_1 + \dots + x_n)}{a n} = \frac{x_1 + \dots + x_n}{n}.$$

Следовательно, при равенстве всех весов можно вычислять вместо взвешенной простую среднюю арифметическую, что значительно сокращает расчеты.

Б. В качестве весов средней вместо абсолютных показателей можно использовать их удельные веса в общем итоге (доли или долины к итогу).

Действительно, удельные веса получаются путем деления абсолютной величины веса каждого варианта на одно и то же число — сумму весов. Следовательно, при использовании удельных весов вместо абсолютных показателей величина средней не изменится, расчеты же ее при этом упрощаются. Кроме того, если показатели, являющиеся весами, известны только в процентах к итогу, использование этих удельных весов — единственный путь, дающий возможность исчислить среднюю.

Допустим, что нужно определить среднюю месячную заработную плату рабочих, если имеются следующие данные.

Таблица 64

ИСПОЛЬЗОВАНИЕ УДЕЛЬНЫХ ВЕСОВ ПРИ РАСЧЕТЕ СРЕДНЕЙ АРИФМЕТИЧЕСКОЙ ВЗВЕШЕННОЙ

Виды и системы оплаты труда	Исходные данные		Расчет числа, пропорционального фонду заработной платы $m_i \bar{x}_i$
	число рабочих (в процентах к итогу) m_i	средняя месячная заработная плата (руб.) \bar{x}_i	
Постоянная основная оплата	42	115	4 830
Временная основная оплата	17	120	2 040
Постоянная повременная оплата	4	104	416
Временная повременная оплата	37	112	4 144
Итого	100	114,3	11 430

Исходной базой расчета средней заработной платы является отношение фонда заработной платы к численности рабочих:

$$\text{Средняя заработная плата одного рабочего} = \frac{\text{Фонд заработной платы}}{\text{Число рабочих}}.$$

Чтобы определить среднюю заработную плату непосредственно на основе этого отношения, т. е. путем деления фонда заработной платы на число рабочих, необходимо знать абсолютные размеры

показателей. Однако данные об абсолютной численности рабочих отсутствуют, а это в свою очередь не дает возможности определить абсолютную величину фонда заработной платы путем умножения средней заработной платы каждой группы рабочих (\bar{x}_i) на их количество.

Единственно возможным способом расчета средней заработной платы является в данном случае использование формулы средней арифметической взвешенной (формула 6.3 — общая средняя из частных средних) и применение в качестве m_i удельных весов отдельных групп рабочих в общей их численности.

Найдя произведение вариантов на веса (см. последнюю графу табл. 6.6) по формуле (6.3) определим среднюю заработную плату:

$$\bar{x} = \frac{\sum \bar{x}_i m_i}{\sum m_i} = \frac{11430}{100} = 114,3 \text{ (руб.)}$$

Если удельные веса выражены не в процентах, а в долях итога, то их сумма равна единице. Формула взвешенной средней арифметической при использовании весов, выраженных в долях, приобретает следующий вид:

$$\bar{x} = \sum x w, \quad (6.6)$$

где $w = \frac{m}{\sum m}$ — удельные веса, выраженные в долях.

Из этой формулы видно, что величина средней зависит не от абсолютного размера весов, а от их относительной величины, от удельного веса в общем итоге. Изменение удельных весов, т. е. изменение структуры (так называемые структурные сдвиги), может привести к изменению средней величины даже при неизменных значениях каждого варианта. В частности, если удельный вес большего варианта увеличится за счет уменьшения удельного веса меньшего варианта, то при прочих равных условиях увеличится и средняя. Наоборот, если удельный вес большего варианта уменьшится за счет увеличения удельного веса меньшего варианта, то соответственно уменьшится и средняя. Покажем это на следующем примере.

Таблица 6.7

К РАСЧЕТУ ВЛИЯНИЯ СООТНОШЕНИЯ ВЕСОВ НА ВЕЛИЧИНУ СРЕДНЕЙ

Период продажи	Исходные данные			Расчет	
	цена (руб. за 1 т) x	количество проданной моркови столовой (в долях итога по колхозу)		$x w'$	$x w''$
		колхоз 1 w'	колхоз 2 w''		
До 15 сентября	60	0,3	0,7	18	42
С 16 сентября	50	0,7	0,3	35	15
Итого	—	1,0	1,0	53	57

Определим по каждому колхозу среднюю цену 1 т моркови:

$$\text{Средняя цена за 1 т моркови} = \frac{\text{Сумма выручки за проданную морковь}}{\text{Количество тонн моркови}}$$

В исходных данных отсутствуют сведения о сумме выручки за проданную морковь. Этот показатель находится в числителе искомого отношения и может быть выражен как сумма произведений цены за 1 т моркови на количество тонн моркови, т. е. на знаменатель исходного отношения. Следовательно, среднюю цену необходимо исчислить по формуле средней арифметической, приняв в качестве весов количество проданной моркови до 15 сентября и с 16 сентября (в нашем примере — удельные веса этих периодов в общем итоге). Расчет средней произведем по формуле 6.3 (см. две последние графы).

Хотя цены, по которым продавалась морковь, одинаковы для обоих колхозов, все же в колхозе 2 средняя цена за 1 т составляет 57 руб., а в колхозе 1 — только 53 руб. Это объясняется тем, что в колхозе 2 в общем количестве проданной моркови выше удельный вес продаж в период до 15 сентября, когда цена за 1 т составляет 60 руб. Разница в величине средних вызвана здесь несомненно различным в каждом колхозе соотношением количества моркови, проданной до и после 15 сентября, т. е. структурными различиями.

Упрощение техники вычисления средней арифметической достигается также путем различных преобразований вариантов искомого признака. При этом используется следующее свойство средней арифметической.

Если все варианты признака увеличить или уменьшить на одно и то же число или в одно и то же число раз, то так же соответственно изменится и средняя.

Пусть

$$\bar{x} = \frac{\sum x m}{\sum m} = \frac{x_1 m_1 + x_2 m_2 + \dots + x_n m_n}{m_1 + m_2 + \dots + m_n}$$

Тогда, увеличив или уменьшив все варианты на одно и то же число a и вычислив среднюю из новых вариантов ($x \pm a$), получим:

$$\overline{(x \pm a)} = \frac{\sum (x \pm a) m}{\sum m} = \frac{\sum x m}{\sum m} \pm a = \bar{x} \pm a.$$

Увеличив все варианты в a раз и вычислив среднюю из новых вариантов (ax), будем иметь:

$$\overline{(ax)} = \frac{\sum ax m}{\sum m} = a \frac{\sum x m}{\sum m} = a \bar{x}.$$

Допустим теперь, что все варианты x сначала уменьшены на одно и то же число x_0 , а затем уменьшены в n раз. Тогда новые



варианты будут выражаться в виде $x' = \frac{x - x_0}{h}$, а их средняя величина будет равна:

$$\bar{x'} = \frac{\sum x'm}{\sum m} = \frac{\sum \left(\frac{x - x_0}{h} \right) m}{\sum m}. \quad (6.7)$$

В то же время согласно рассмотренному свойству средняя из новых вариантов $\left(\frac{x - x_0}{h} \right)$ будет равна средней из первоначальных вариантов, уменьшенной соответственно сначала на x_0 , а затем в h раз:

$$\bar{x'} = \frac{\bar{x} - x_0}{h},$$

$$\bar{x} = x_0 + h\bar{x'}. \quad (6.8)$$

откуда

Формулы (6.7) и (6.8) в ряде случаев позволяют значительно упрощать расчеты средней. Величины x_0 и h являются при этом произвольными, однако наибольшее упрощение расчетов достигается в тех случаях, когда в качестве x_0 принимается один из центральных вариантов ряда (или середина одного из центральных интервалов), а в качестве h — общий наибольший делитель величин $(x - x_0)$. При равных интервалах в качестве h наиболее целесообразно использовать величину интервала.

Величина x_0 носит название начала отсчета или условного нуля. Отсюда и сам упрощенный способ расчета называется способом отсчета от условного начала (условного нуля). Средняя из отклонений вариантов от некоторой постоянной величины x_0 называется в математической статистике первым моментом, в связи с чем и сам способ называется также способом моментов.

Значительное упрощение вычисления средней способом моментов достигается для вариационных рядов с равными интервалами. Например, нужно определить среднюю себестоимость 1 т молока в колхозах по следующим данным (см. табл. 6.8).

Исходной базой расчета средней себестоимости 1 т молока является отношение

$$\frac{\text{Общая себестоимость всего молока}}{\text{Количество молока в тоннах}}.$$

Обозначив себестоимость 1 т молока через x , а количество молока через m , найдем, что

$$\bar{x} = \frac{\sum xm}{\sum m}.$$

Расчет произведем по формуле (6.8). За начало условного отсчета примем середину третьего интервала: $x_0 = 122,5$, а в качестве h — величину интервала ($h = 5$). Отнимая от середины каждого

Таблица 6.8

распределение колхозов по себестоимости 1 т молока
Упрощенный способ расчета средней арифметической по данным
равномерного ряда с равными интервалами)

Исходные данные			Расчет		
Варианты себестоимости 1 т молока (руб.)	число колхозов, находящих данную себестоимость (в процентах к итогу)	в них произ- ведено молока (в процентах к итогу) m	средняя интервала x	$x' = \frac{x - x_0}{h}$	$\left(\frac{x' m}{n} \right) \times m$
100—115	7	9	112,5	-2	-18
115—130	16	18	117,5	-1	-18
130—145	25	24	122,5	0	0
145—160	28	27	127,5	+1	+27
160—175	18	18	132,5	+2	+36
Итого	5	4	137,5	+3	+12
Итого	100	100	—	—	+75 -36 +39

интервала 122,5 и деля полученные разности на 5, получим новые варианты $x' = \frac{x - x_0}{h} = \frac{x - 122,5}{5}$ (см. гр. 4). Перемножив эти варианты на соответствующие им веса m и сложив полученные произведения (см. гр. 5), по формуле 6.7 найдем среднюю величину новых вариантов:

$$\bar{x'} = \frac{\sum x'm}{\sum m} = \frac{39}{100} = 0,39.$$

Следовательно, $\bar{x} = x_0 + h\bar{x'} = 122,5 + 5 \times 0,39 = 124,45$, т. е. средняя себестоимость 1 т молока составляет 124 р. 45 к.

Отметим, что при равных интервалах нет необходимости находить ни середину каждого интервала, ни величины $(x - x_0)$, так как новые варианты $x' = \frac{x - x_0}{h}$ могут быть получены чисто механически, без специальных расчетов. Действительно, при равных интервалах как середины соседних интервалов (см. гр. 3), так и разности $(x - x_0)$ отличаются друг от друга ровно на величину интервала. Следовательно, после деления этих разностей на величину интервала значения соседних между собой новых вариантов будут отличаться друг от друга на единицу. Поэтому для получения новых вариантов (см. гр. 4) достаточно против одного из центральных интервалов поставить нуль и механически писать вверх и вниз ряд натуральных чисел со знаком минус, а вниз — со знаком плюс. При этом условным началом отсчета (x_0) будет середина этого интервала, против которого в графе x' написан нуль, а h будет равно величине интервала.

Средняя гармоническая

В ряде случаев исходная база расчета средней и имеющиеся для этого расчета данные приводят не к средней арифметической, а к другой форме средней — к средней гармонической.

Например, необходимо определить средний годовой удой молока от одной коровы в целом по колхозам и совхозам области по следующим данным.

Таблица 69

К РАСЧЕТУ СРЕДНЕЙ ГАРМОНИЧЕСКОЙ ВЗВЕШЕННОЙ

	Исходные данные		Расчет
	средний годовой удой молока от одной коровы (кг) x	общее производство молока (тыс. т) M	среднее поголовье коров (тыс. голов) $\frac{M}{x}$
Колхозы	2400	360	150
Совхозы	3000	300	100
Итого	2640	660	250

Исходная база расчета среднего удоя молока от одной коровы выражается так:

$$\text{Средний годовой удой молока от одной коровы} = \frac{\text{Валовой надой молока за год от всех коров (общее производство молока)}}{\text{Среднее поголовье коров за год}}$$

Величина валового надоя молока, т. е. показателя, который находится в числителе исходного отношения, известна. Чтобы определить неизвестную величину среднего поголовья коров, нужно (отдельно по колхозам и совхозам) разделить валовой надой молока на средний удой от одной коровы. Так, разделив 360 тыс. т на 2,4 т, найдем, что среднее поголовье коров в колхозах составляло 150 тыс. голов. Точно так же определим, что среднее поголовье коров в совхозах составляло 100 тыс. голов.

Тогда в соответствии с исходной базой средний удой от одной коровы в целом по колхозам и совхозам составит:

$$\bar{x} = \frac{660 \text{ тыс. т}}{250 \text{ тыс. голов}} = 2640 \text{ кг.}$$

Обозначив варианты среднего удоя через x , а валовой надой (общее производство) молока через M , действия по расчету среднего удоя можно выразить в виде следующей формулы:

$$\bar{x} = \frac{M_1 + M_2 + \dots + M_n}{\frac{M_1}{x_1} + \frac{M_2}{x_2} + \dots + \frac{M_n}{x_n}} = \frac{\sum M}{\sum \frac{M}{x}} \quad (6.9)$$

Это — формула средней гармонической взвешенной, где M — веса.

Таким образом, чтобы исчислить взвешенную среднюю гармоническую, нужно сумму весов разделить на сумму результатов деления каждого веса на соответствующий вариант признака.

Если вес каждого варианта равен единице, то при n вариантах формула средней гармонической получает такой вид:

$$\bar{x} = \frac{n}{\sum \frac{1}{x}} \quad (6.10)$$

Это — простая средняя гармоническая.

Формулу средней гармонической можно получить также из уравнения средней.

Взяв в нашем примере (см. табл. 69) поголовье коров по каждой категории хозяйства выразить как частное от деления валового надоя молока на удой от одной коровы $\left(\frac{M}{x}\right)$, а в целом по колхозам и совхозам — как $\frac{M_1}{x_1} + \frac{M_2}{x_2}$. Умножив варианты (x) средней величиной (\bar{x}) и приравняв новую сумму к старой, получим:

$$\frac{M_1}{x_1} \times \bar{x} = \frac{M_1}{x_1} \times \frac{M_2}{x_2} = \frac{M_1}{x_1} + \frac{M_2}{x_2} = \frac{M_1 + M_2}{x_1 + x_2}$$

Откуда

$$\bar{x} = \frac{M_1 + M_2}{\frac{M_1}{x_1} + \frac{M_2}{x_2}}$$

или в общем виде

$$\bar{x} = \frac{\sum M}{\sum \frac{M}{x}}$$

В рассмотренном примере, исчисляя среднюю, мы располагаем наряду с вариантами осредняемого признака слагаемыми (M) этого показателя, который находится в числителе исходного отношения. Эти слагаемые числителя использовались в качестве весов. Величина же показателя, который находится в знаменателе исходного отношения, непосредственно не была известна. Следовательно, взвешенная средняя гармоническая применяется в тех случаях, когда показатель, величина которого непосредственно неизвестна, находится в знаменателе исходного отношения, а в качестве весов используются слагаемые показателя, находящегося в числителе этого отношения.

Величина средней гармонической, так же как и величина средней арифметической, не изменяется при увеличении или уменьшении всех весов в одно и то же число раз. Это дает возможность использовать в качестве весов их проценты, а в частности удельные веса — проценты к итогу.

Например, требуется определить средний процент выполнения плана по следующим данным (см. табл. 6.10).

$$\text{Процент выполнения плана} = \frac{\text{Фактический уровень}}{\text{Планировый уровень}} \times 100.$$

ИСПОЛЬЗОВАНИЕ УДЕЛЬНЫХ ВЕСОВ КАК ВЕСОВ
СРЕДНЕЙ ГАРМОНИЧЕСКОЙ

Заводы	Исходные данные		Расчет числа, пропорционального выпуску продукции по плану $\frac{M}{x}$
	процент выполнения плана по выпуску продукции x	фактический выпуск продукции (в процентах к итогу) M	
1	98	39,8	0,406 0,557
2	108	60,2	
Итого	...	100,0	0,963

В данном случае ни фактический, ни плановый выпуск продукции в абсолютном выражении не известны и не могут быть найдены. Поэтому найти средний процент выполнения плана на базе одного исходного отношения нельзя.

В исходных данных нет сведений о выпуске продукции по плану (знаменатель исходного отношения), в качестве же весов могут быть использованы удельные веса предприятий в общем итоге фактического выпуска продукции (числитель исходного отношения). Следовательно, налицо условия применения взвешенной средней гармонической:

$$\bar{x} = \frac{\sum M}{\sum \frac{M}{x}} = \frac{100}{\frac{39,8}{98} + \frac{60,2}{108}} = \frac{100}{0,406 + 0,557} = \frac{100}{0,963} = 103,8\%.$$

3. МЕДИАНА И МОДА

Особого рода средними, используемыми в экономическом анализе для изучения структуры вариационного ряда, являются медиана и мода.

Медиана

Медианой в статистике называется значение признака у той единицы совокупности, которая расположена в середине упорядоченного ряда. Если, например, выстроить по росту 9 студентов, то рост пятого по счету от начала (или от конца) ряда будет являться медианой. Если же студентов будет 10, то медианой будет среднее значение роста пятого и шестого студентов.

В вариационном ряду медианой является значение признака у той единицы совокупности, которая делит вариационный ряд по сумме частот на две равные части, так что у половины единиц значение признака меньше медианы, а у другой половины — больше ее.

Медиану обычно определяют по данным интервального вариационного ряда. Это связано с тем, что при небольшом числе дискретных вариантов значение признака, равное медиане, могут иметь очень многие единицы совокупности, так что медиана становится непоказательной.

Таблица 6.10

Работники, рабочие промышленного предприятия так распределяются по квалификации (разрядности) на конец года:

Таблица 6.11

РАСПРЕДЕЛЕНИЕ МЕДИАНЫ В ДИСКРЕТНОМ ВАРИАЦИОННОМ РЯДУ

Разряд	Количество рабочих	Количество рабочих, имеющих данный или меньший разряд (ряд накопленных частот)
2-й	44	44
3-й	255	44 + 255 = 299
4-й	490	299 + 490 = 789
5-й	370	789 + 370 = 1 159
6-й	120	1 159 + 120 = 1 279
Итого	1 279	—

Чтобы найти порядковый номер рабочего, разряд которого является медианой, нужно к общему числу рабочих прибавить единицу и сумму поделить на 2: $1279 + 1 : 2 = 640$. Как видно из ряда накопленных частот, рабочий, имеющий этот порядковый номер, имеет 4-й разряд. Следовательно, медиана равна 4. Однако 4-й разряд имеют 490 рабочих. При этом меньше четвертого разряда имеют 299 рабочих, а больше 4-го разряда — 490 человек.

Определение медианы для интервального ряда рассмотрим на примере.

Таблица 6.12

РАСПРЕДЕЛЕНИЕ ЧИСЛА МУЖЧИН, ВСТУПИВШИХ В БРАК
в 1973 г. ПО ВОЗРАСТУ (СССР)

Возраст при вступлении в брак	Число мужчин в процентах к итогу (частоты)	Накопленный ряд частот (процент мужчин, вступивших в брак в возрасте, меньший верхней границы интервала)
До 20	4,3	4,3
20—25	57,5	61,8
25—30	17,1	78,9
30—35	7,6	86,5
35—40	4,3	90,8
40—50	4,3	95,1
50 и старше	4,9	100,0
Итого	100	—

В данном случае медианой является тот возраст, моложе которого вступили в брак 50% мужчин. Из данных последней графы таблицы видно, что в возрасте моложе 20 лет вступили в брак лишь 4,3% мужчин, а в возрасте моложе 25 лет — 61,8%, т. е. уже более 50% всех мужчин. Следовательно, медиана, которой соответствует накопленная частота 50%, лежит где-то между 20 и 25 годами, т. е. во втором возрастном интервале, который является медианным.

Накопленной частоты I интервала до 50% не хватает 45,7% (50%—4,3%). Следовательно, чтобы выделить половину мужчин от начала ряда к накопленной частоте I интервала нужно добавить еще 4,3%, взятых из частоты медианного (II) интервала. Эти 45,7% составляют 0,795 (45,7 : 57,5) общей частоты этого интервала. Предположим далее, что мужчины, которым при вступлении в брак было от 20 до 25 лет, распределяются по возрасту приблизительно равномерно. Тогда 0,795 частоты медианного интервала будет соответствовать 0,795 величины этого интервала, равная 4,0 (0,795 × 5) года. Таким образом, если накопленная частота для верхней границы I интервала (20 лет) меньше 50% на 45,7%, то верхняя граница в свою очередь на 4,0 года меньше медианы, которая равна 20+4,0=24 года.

Действия, произведенные при расчете медианы, могут быть обобщены в виде следующей формулы:

$$Me = x_{Me} + h_{Me} \frac{\frac{\sum m}{2} - S_{Me-1}}{m_{Me}}, \quad (6.11)$$

где x_{Me} — начало (нижняя граница) медианного интервала;

h_{Me} — величина медианного интервала;

m_{Me} — частота медианного интервала;

S_{Me-1} — частота, накопленная до медианного интервала.

Подставив в эту формулу данные нашего примера, получим:

$$Me = 20 + 5 \times \frac{50 - 4,3}{57,5} = 24,0 \text{ (года)},$$

т. е. половина вступивших в брак мужчин были моложе 24 лет, а половина — старше этого возраста.

Применение медианы вместо средней или наряду с ней целесообразно при наличии в вариационном ряду открытых интервалов. Для исчисления медианы обычно не требуется условного установления неизвестных границ открытых крайних интервалов, так что отсутствие сведений о них не влияет на точность расчета медианы.

Целесообразно применение медианы и в тех случаях, когда показатели, которые должны быть приняты в качестве весов средней, неизвестны. Медиану применяют вместо средней арифметической при статистических методах контроля качества продукции.

Сумма абсолютных значений отклонений вариантов от медианы меньше, чем от любого другого числа:

$$\sum |x - Me| = \min.$$

Это свойство медианы может быть использовано для определения места сооружения зданий и пунктов массового посещения и использования (школ, детских садов, яслей и т. п., а также спортивных пунктов, уличных водоразборных колонок, телефонных ка-

бин и др.) Так, если на улице поселка предполагается поставить лавочку, телефона-автомата, то ее целесообразно расположить в такой точке, которая делит пополам не длину улицы, а число жителей, т. е. с таким расчетом, чтобы по обе стороны от лавочки проживало примерно равное количество жителей.

При изучении структуры вариационного ряда используются также квадратичный ряд (по сумме частот) на 4 равные части, и децилы, которые делят ряд на 10 равных частей. Формулы их строятся аналогично (6.11).

Мода

Модой в статистике называется наиболее часто встречающееся значение признака. Применительно к вариационному ряду мода — это вариант, который имеет наибольшую частоту, т. е. повторяется наибольшее число раз.

Мода имеет большое практическое применение. Например, в легкой промышленности при конструировании колодок для массового пошива обуви необходимо установить ряд параметров: соотношение ширины ступни и ее длины, соотношение высоты подъема и длины ступни и др. При установлении этих параметров мода имеет явные преимущества перед средней.

Известно, что соотношение высоты подъема и длины ступни неодинаково. При индивидуальном пошиве обуви это обстоятельство учитывается в каждом конкретном случае. При массовом пошиве обуви нужен один обобщающий показатель. Если высоту подъема для каждого размера обуви определить как среднюю арифметическую, то может случиться, что для большинства людей обувь будет либо слишком тесной, либо слишком свободной в подъеме. Если же для каждого размера обуви принять модальное, т. е. наиболее часто встречающееся соотношение, то для большинства людей обувь, сшитая по такой выкройке, окажется удобной.

При изучении цен колхозного рынка регистрируют цены, которые преобладают на рынке, т. е. модальные цены. В подобных случаях мода является значительно лучшей обобщающей характеристикой, чем средняя арифметическая.

В ряде случаев только мода может дать обобщающую характеристику общественных явлений.

Например, чтобы определить 1-е место среди соревнующихся бригад цеха по результатам работы за год берут данные о том, сколько раз в течение года присуждалось 1-е место каждой из них. Та бригада, которая чаще других выходила на 1-е место в году, и получает звание бригады-победительницы, занявшей 1-е место по результатам работы за год.

В спортивных соревнованиях при присуждении участникам соревнований мест принимается во внимание место, названное большинством судей. Так, например, присуждаются места на европейских турнирах танцев.

При определении моды в дискретном вариационном ряду не требуется никаких формул: модой будет тот вариант ряда, который имеет наибольшую частоту, т. е. чаще других повторяется.

Так, в распределении рабочих по квалификации (см. табл. 6.11) мода равна 4.

В непрерывных вариационных рядах с равными интервалами мода определяется по следующей приближенной формуле:

$$M_o = x_{M_o} + h_{M_o} \cdot \frac{m_{M_o} - m_{M_o-1}}{(m_{M_o} - m_{M_o-1}) + (m_{M_o} - m_{M_o+1})}, \quad (6.12)$$

где x_{M_o} — начало (нижняя граница) модального интервала, т. е. интервала, имеющего наибольшую частоту;
 h_{M_o} — величина модального интервала;
 m_{M_o} — частота (или частость) модального интервала;
 m_{M_o-1} — частота (или частость) предмодального интервала;
 m_{M_o+1} — частота (или частость) постмодального интервала.
 Покажем применение этой формулы на следующем примере.

Таблица 6.13
 РАСПРЕДЕЛЕНИЕ ЧИСЛЕННОСТИ РАБОЧИХ И СЛУЖАЩИХ
 НЕФТЕПЕРЕРАБАТЫВАЮЩЕЙ ПРОМЫШЛЕННОСТИ
 СССР ПО ОБЩЕМУ СТАЖУ РАБОТЫ
 (на 1 июня 1967 г.)

Общий стаж работы (лет)	Численность рабочих и служащих (процентов)
0—5	16,4
5—10	22,2
10—15	22,6
15—20	18,5
20—25	11,3
25 и выше	9,0
Итого	100

$$M_o = 10 + 5 \times \frac{22,6 - 22,2}{(22,6 - 22,2) + (22,2 - 18,5)} = 10,4 \text{ (лет)}.$$

Следовательно, в 1967 г. у рабочих нефтеперерабатывающей промышленности СССР относительно чаще встречался общий производственный стаж — 10,4 года.

Если частоты предмодального и постмодального интервалов равны между собой, то мода равна середине модального интервала. Если частота постмодального интервала больше, чем частота предмодального интервала, то мода больше середины модального интервала; наоборот, если постмодальная частота меньше предмодальной, мода тоже меньше середины модального интервала, как это имеет место в нашем примере.

ПОКАЗАТЕЛИ ВАРИАЦИИ

Неизменность признаков вариации

Средняя величина представляет обобщающую характеристику варьирующего признака. Однако, характеризуя вариационный ряд в целом, средняя не показывает, как распределяются около нее варианты осредняемого признака, сосредоточены ли они вблизи средней или значительно отклоняются от нее. Средняя не показывает также характер вариации признака и величину его колеблемости. Это можно показать на таком примере.

Предположим, что одинаковую работу выполняют две бригады, каждая из трех человек. Пусть количество деталей, изготовленных за смену отдельными рабочими, составляло в первой бригаде 95, 101 и 105 шт., а во второй бригаде — 75, 100 и 125 шт. Средняя выработка на одного рабочего в обеих бригадах одинакова и составляет 100 шт., однако колеблемость выработки отдельных рабочих около средней выработки в первой бригаде значительно меньше, чем во второй. Очевидно, что этому причиной ряд факторов: квалификация рабочих, их производственный стаж, возраст, интенсивность труда и т. д.

Воздействием множества факторов вызвана вариация производительности труда, заработной платы, себестоимости продукции, прибыли, рентабельности и т. д. Изучая количественную сторону случайных явлений, статистика почти всегда имеет дело с варьированием признаков изучаемых ею явлений. Задача статистики — дать числовое выражение колеблемости признаков для более глубокого и всестороннего познания сущности объектов ее изучения. Для этого используются различные показатели вариации.

Размах вариации

Наиболее простым показателем вариации является ее *размах*, представляющий разность между наибольшим и наименьшим значениями признака:

$$\text{размах вариации} = x_{\max} - x_{\min}. \quad (6.13)$$

В приведенном примере размах вариации сменной выработки составляет для первой бригады 10 деталей (105—95), а для второй — 50 деталей (125—75), т. е. в 5 раз больше. Это позволяет расширить наши представления о результатах производственной деятельности бригад. Запишем для каждой бригады не только среднюю выработку, но и размах вариации осредненных признаков:

$$\text{первая бригада} - \bar{x}_1 = 100(95-105),$$

$$\text{вторая бригада} - \bar{x}_2 = 100(75-125).$$

Становится очевидным, что хотя средняя сменная выработка у обеих бригад численно одинакова, однако качество этой средней неодинаково. Средняя выработка в первой бригаде более типична, чем во второй, а средняя выработка второй бригады менее «устой-

чивая, чем в первой, в смысле возможного изменения ее размеров.

Другая сторона рассматриваемого вопроса заключается в определении возможных резервов роста выработки. Средняя выработка не может служить базой расчета возможных резервов роста выработки. Такой базой может служить размах вариации. Так, можно сказать, что если в первой бригаде все рабочие достигнут максимальной в данной бригаде выработки, то всего будет изготовлено 315 шт. деталей (3×105), если во второй бригаде все рабочие достигнут максимальной выработки, то всего будет изготовлено 375 шт. деталей (3×125). Резерв роста выработки первой бригады составляет 15 шт., второй — 75 шт. Вторая бригада обладает большими резервами роста выработки, чем первая.

Размах вариации выражается в тех именованных числах, в каких выражены варианты. К числу достоинств данного показателя вариации относится простота его исчисления. Вместе с тем ему присущи и некоторые недостатки.

Особенность показателя размаха вариации заключается в том, что он зависит лишь от двух крайних значений признака. По этой причине его целесообразно применять в тех случаях, когда особое значение имеет либо минимальный, либо максимальный вариант (например, при испытании стальных тросов на разрыв и т. п.). Другая сторона этой особенности заключается в том, что на величину размаха вариации большое влияние оказывает случайность. Так как из статистического ряда берутся только два значения признака, причем крайние в ряду, на размер которых могут оказывать влияние причины случайного характера, то и размах вариации может быть записанном от причин случайного характера.

С отмеченной особенностью связано и то обстоятельство, что показатель размаха вариации не учитывает частот в вариационном ряду.

В силу этих недостатков статистикой разработаны более точные показатели вариации и применение показателя размаха вариации имеет ограниченный характер. Однако как первая характеристика вариационного ряда показатель размаха вариации незаменим. Размах вариации может быть использован в ряде случаев для сравнительной характеристики варьирования признака только по вариантам, без учета частот.

Например, размах вариации возраста у студентов дневных отделений вузов обычно меньше, чем у студентов вечерних отделений, размах вариации возраста невест, вышедших замуж за мужчин моложе 20 лет, обычно меньше, чем размах вариации возраста женихов, женившихся на женщинах моложе 20 лет, и т. д.

Стремление составить показатель вариации, который опирался бы на все значения признака, приводит к среднему линейному отклонению. Среднее линейное отклонение представляет среднюю арифметическую из абсолютных значений отклонений отдельных вариантов от их средней арифметической.

Если каждый вариант повторяется один раз, то среднее линейное отклонение (Л) равно:

$$Л = \frac{\sum |x - \bar{x}|}{n}, \quad (6.14)$$

где $|x - \bar{x}|$ — абсолютные значения отклонений отдельных вариантов от их средней арифметической;
 n — число членов ряда.

Для вариационного ряда с неравными частотами формула имеет такой вид:

$$Л = \frac{\sum |x - \bar{x}| m}{\sum m}, \quad (6.15)$$

где $\sum m$ — сумма частот вариационного ряда.

Прямые скобки у отклонений указывают, что эти отклонения при суммировании или умножении берутся без учета их знаков, т. е. в виде абсолютных значений. Делается это потому, что, как известно из предыдущего (см. с. 162), алгебраическая сумма отклонений вариантов от средней арифметической равна нулю.

Для иллюстрации расчета среднего линейного отклонения допустим, что у нас имеются следующие данные о распределении рабочих двух предприятий по квалификации (тарифным разрядам) по состоянию на начало отчетного года:

Таблица 6.14

РАСПРЕДЕЛЕНИЕ РАБОЧИХ ПО ТАРИФНЫМ РАЗРЯДАМ НА НАЧАЛО ГОДА

Квалификационный разряд	число рабочих (в процентах к итогу по предприятиям)		Расчет среднего линейного отклонения					
			предприятие № 1			предприятие № 2		
	предприятие № 1	предприятие № 2	$x m_1$	$x - \bar{x}_1$	$ x - \bar{x}_1 m_1$	$x m_2$	$x - \bar{x}_2$	$ x - \bar{x}_2 m_2$
	m_1	m_2						
1	9	17	9	-2	18	17	-2	34
2	21	19	42	-1	21	38	-1	19
3	40	31	120	0	0	93	0	0
4	23	18	92	1	23	72	1	18
5	5	10	25	2	10	50	2	20
6	2	5	12	3	6	30	3	15
Итого	100	100	300	—	78	300	—	106

Средний разряд рабочих предприятия № 1 равен $\bar{x}_1 = 300 : 100 = 3$. Такой же разряд имеют в среднем и рабочие предприятия № 2. Однако колеблемость этого признака по предприятию № 2 значительно больше:

$$Л_1 = 78 : 100 = 0,78; \quad Л_2 = 106 : 100 = 1,06.$$

СРЕДНЕЕ КВАДРАТИЧЕСКОЕ ОТКЛОНЕНИЕ

Типовой вариант i	Число рабочих (в процентах к итогу по предприятию)		$x - \bar{x}_1 =$ $= x - \bar{x}_2$	$(x - \bar{x}_1)^2 =$ $= (x - \bar{x}_2)^2$	$(x - \bar{x}_1)^2 m_1$ $(x - \bar{x}_2)^2 m_2$	$(x - \bar{x}_1)^2 m$ $(x - \bar{x}_2)^2 m$
	предприя- тие № 1 m_1	предприя- тие № 2 m_2				
A	1	2	3	4	5	6
1	9	17	-2	4	36	68
2	21	19	-1	1	21	19
3	40	31	0	0	0	0
4	23	18	1	1	23	18
5	5	10	2	4	20	40
6	2	5	3	9	18	45
Итого	100	100	—	—	118	190

Это свидетельствует о том, что состав рабочих по квалификации на предприятии № 1 более однороден, чем на предприятии № 2.

Среднее линейное отклонение обладает значительным преимуществом перед размахом вариации в отношении полноты характеристики колеблемости признака. Однако при этом нарушается элементарное правило математики, так как отклонения от среднего значения признака складываются без учета знаков. Это побуждает искать такой показатель вариации, который был бы лишен и этого недостатка.

Отмеченный выше недостаток показателя среднего линейного отклонения может быть устранен путем возведения в квадрат отклонений вариантов от средней арифметической. Таким путем мы приходим к новому показателю вариации — дисперсии. Дисперсией (обозначается она греческой буквой сигма малая в квадрате — σ^2) называется **средний квадрат отклонений вариантов от их средней величины**. Порядок ее исчисления можно выразить следующей формулой:

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 m}{\sum m}, \quad (6.16)$$

а при равенстве весов или их отсутствии соответственно:

$$\sigma^2 = \frac{\sum (x - \bar{x})^2}{n}. \quad (6.17)$$

Среднее квадратическое отклонение представляет собой **корень квадратный из дисперсии**. Общая формула этого показателя может быть записана так:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 m}{\sum m}}, \quad (6.18)$$

либо при равенстве весов или их отсутствии соответственно:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}. \quad (6.19)$$

Среднее квадратическое отклонение измеряет абсолютный размер колеблемости признака и выражается в тех же единицах измерения, что и варианты.

Порядок расчета дисперсии и среднего квадратического отклонения рассмотрим на том же примере распределения рабочих двух предприятий по разрядам, используя уже исчисленные средние ($x_1 = x_2 = 3$; см. табл. 6.15).

Так как варианты и средние по обоим предприятиям одинаковы, отклонения от средней (гр. 3) и квадраты этих отклонений

(гр. 4) тоже совпадают. Подставляя итоги гр. 5 и 6 в формулу (6.16) получим значения дисперсий:

$$\sigma_1^2 = \frac{118}{100} = 1,18 \quad \text{и} \quad \sigma_2^2 = \frac{190}{100} = 1,90,$$

извлекая теперь корень квадратный, имеем: $\sigma_1 = \sqrt{1,18} = 1,086$ (разряда); $\sigma_2 = \sqrt{1,90} = 1,378$ (разряда).

Исчисление дисперсий по формуле (6.16) часто является очень трудоемким. Для упрощения расчетов применяется ряд способов.

В тех случаях, когда варианты выражаются небольшими и незначительными числами, дисперсия может быть найдена по следующей формуле:

$$\sigma^2 = \bar{x}^2 - \bar{x}^2, \quad (6.20)$$

$$\bar{x}^2 = \frac{\sum x^2 m}{\sum m}, \quad \text{а} \quad \bar{x}^2 = \left(\frac{\sum x m}{\sum m} \right)^2.$$

Формула (6.20) получается путем преобразования формулы (6.16), в которой сначала производится возведение в квадрат и сложение на m , а затем почленное деление и сложение.

Формула (6.20) читается так: **дисперсия равна разности между средним квадратов вариантов и квадратом их средней**.

Например, на станках-автоматах обработано 640 деталей, причем на каждую деталь затрачивалось по 0,5 часа, а на обычных станках обработано 360 деталей при трудоемкости одной детали

1 час. Найдем дисперсию затрат труда на одну деталь, используя формулу (6.20):

$$\bar{x} = \frac{\sum xm}{\sum m} = \frac{0,5 \times 640 + 1 \times 360}{1000} = 0,68 \text{ (часа);}$$

$$\overline{x^2} = \frac{\sum x^2 m}{\sum m} = \frac{(0,5)^2 \times 640 + 1^2 \times 360}{1000} = \frac{160 + 360}{1000} = 0,52;$$

$$\sigma^2 = \overline{x^2} - \bar{x}^2 = 0,52 - (0,68)^2 = 0,52 - 0,46 = 0,06$$

и среднее квадратическое отклонение $\sigma = \sqrt{0,06} = 0,24$ (часа).

Второй (упрощенный) способ основан на отсчете от условного начала (способ моментов). Формула дисперсии при расчете по этому способу такова:

$$\sigma^2 = h^2 [(\overline{x'})^2 - (\bar{x}')^2], \quad (6.21)$$

где

$$x' = \frac{x - x_0}{h}; \quad \bar{x}' = \frac{\sum x' m}{\sum m} = \frac{\sum \left(\frac{x - x_0}{h} \right) m}{\sum m};$$

$$(\overline{x'})^2 = \frac{\sum (x')^2 m}{\sum m} = \frac{\sum \left(\frac{x - x_0}{h} \right)^2 m}{\sum m}.$$

Особенно упрощаются расчеты по формуле (6.21) по данным интервального ряда с равными интервалами. Покажем это на приводившемся ранее примере себестоимости 1 т молока.

Таблица 6.18

К РАСЧЕТУ СРЕДНЕГО КВАДРАТИЧЕСКОГО ОТКЛОНЕНИЯ СПОСОБОМ ОТСЧЕТА ОТ УСЛОВНОГО НУЛЯ ПРИ РАВНЫХ ИНТЕРВАЛАХ ВАРИАНЦИОННОГО РЯДА

Группы колозов по уровню себестоимости 1 т молока (руб.)	Произведено молока (в процентах к итогу) m	$x' = \frac{x - x_0}{h}$	$x' m = \left(\frac{x - x_0}{h} \right) m$	$(x')^2 m = \left(\frac{x - x_0}{h} \right)^2 m$
А	1	2	3	гр. 4 - гр. 2 × гр. 3
110-115	9	-2	-18	36
115-120	18	-1	-18	36
120-125	24	0	0	0
125-130	27	1	+27	27
130-135	18	2	+36	72
135 и более	4	3	+12	36
Итого	100	-	+39	189

Преобразованные варианты x' (гр. 2) пишем механически (см. с. 167), что приводит к $x_0 = 122,5$ (середина интервала, против

которого в гр. 2 написан ноль) и $h = 5$ (величина интервала). Изымшие расчеты таковы:

$$\bar{x}' = \frac{\sum x' m}{\sum m} = 0,39; \quad (\bar{x}')^2 = 0,1521;$$

$$(\overline{x'})^2 = \frac{\sum (x')^2 m}{\sum m} = 1,89.$$

$$\sigma = h \sqrt{(\overline{x'})^2 - (\bar{x}')^2} = 5 \sqrt{1,89 - 0,1521} = 5 \times 1,318 = 6,59 \text{ (руб.).}$$

В ряде случаев изучают не среднюю величину признака, а долю единиц, обладающих или не обладающих тем или иным признаком, например доля (или процент) бракованных изделий, процент лиц, имеющих высшее образование, средний вес студентов, получающих стипендию, и т. п. При такой альтернативной вариации, когда имеются лишь два исключаящие друг друга варианта (наличие и отсутствие признака у единицы), наличие признака обозначается через 1, а его отсутствие — через 0. Обозначим долю единиц, обладающих данным признаком через w , а долю единиц, которые этим признаком не обладают, через $q = 1 - w$, найдем среднюю и дисперсию.

Таблица 6.17

К ВЫВОДУ ФОРМУЛЫ ДИСПЕРСИИ АЛЬТЕРНАТИВНОГО ПРИЗНАКА

	x	m	$x m$	$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 m$
Наличие признака	1	w	w	$1 - w = q$	q^2	$q^2 w$
Отсутствие признака	0	q	0	$-w = -w$	w^2	$w^2 q$
Итого	-	1	w	-	-	$w q$

$$\bar{x} = \frac{\sum xm}{\sum m} = w; \quad (6.22)$$

$$\sigma_w^2 = \frac{\sum (x - \bar{x})^2 m}{\sum m} = q^2 w + w^2 q = w q (q + w).$$

Так как $w + q = 1$, получаем

$$\sigma_w^2 = w q = w (1 - w) \quad (6.23)$$

$$\sigma = \sqrt{w q}.$$

Если, например, 60% студентов имеют производственный стаж, а 40% студентов его не имеют, то среднее квадратическое отклонение для доли будет равно:

$$\sigma_w = \sqrt{0,6 \times 0,4} = \sqrt{0,24}, \text{ откуда } \sigma_w^2 = 0,49, \text{ или } 49\%.$$

Коэффициенты вариации

Среднее линейное и среднее квадратическое отклонения являются мерой абсолютной колеблемости признака и всегда выражаются в тех же единицах измерения, в которых выражен изучаемый признак. Это не позволяет сопоставлять между собой средние отклонения различных показателей. Чтобы иметь возможность производить такие сопоставления, средние отклонения часто выражаются в процентах к средней арифметической т. е. выражаются в виде относительных величин.

Процентное отношение среднего линейного или среднего квадратического отклонения к средней арифметической называется коэффициентом вариации и обозначается V :

$$V = \frac{l}{x} \cdot 100, \quad (6.24)$$

либо

$$V = \frac{\sigma}{x} \cdot 100. \quad (6.25)$$

Так, для приведенного выше примера вариации тарифных разрядов рабочих на двух предприятиях коэффициенты вариации составляют:

а) на основе среднего линейного отклонения

$$V_1 = \frac{0,77 \times 100}{3} = 25,7\% \quad \text{и} \quad V_2 = \frac{1,06 \times 100}{3} = 35,3\%,$$

б) на основе среднего квадратического отклонения

$$V_1 = \frac{1,085 \times 100}{3} = 36,2\% \quad \text{и} \quad V_2 = \frac{1,378 \times 100}{3} = 45,9\%,$$

т. е. среднее линейное отклонение составляет на первом предприятии 25,7% средней арифметической, а на втором — 35,3%, а среднее квадратическое отклонение — соответственно 36,2 и 45,9%.

Являясь относительными величинами, коэффициенты вариации могут сопоставляться не только для одинаковых, одноименных показателей, но и для различных показателей. Например, коэффициент вариации тарифных разрядов можно сравнивать с коэффициентами вариации стажа работы, заработной платы и других показателей того же коллектива рабочих.

Если некоторая совокупность единиц делится на группы, то наряду с общей дисперсией могут быть также найдены дисперсии для каждой отдельной группы — *групповые* (внутригрупповые) *дисперсии* (σ_i^2), а также их средняя величина:

$$\bar{\sigma}^2 = \frac{\sum \sigma_i^2 m_i}{\sum m_i}. \quad (6.26)$$

Кроме того, может быть вычислена также *межгрупповая дисперсия* (δ^2), характеризующая колеблемость групповых (частных) значений (\bar{x}_i) около общей средней (\bar{x}):

$$\delta^2 = \frac{\sum (\bar{x}_i - \bar{x})^2 m_i}{\sum m_i}. \quad (6.27)$$

В математической статистике доказывается, что между общей дисперсией (σ^2), средней из групповых дисперсий и межгрупповой дисперсией существует такая связь:

$$\sigma^2 = \bar{\sigma}^2 + \delta^2. \quad (6.28)$$

Формула (6.28) выражает *правило сложения дисперсий* (важность): *общая дисперсия равна средней из групповых дисперсий плюс межгрупповая дисперсия*.

Например, имеются следующие данные о дневной выработке рабочих 2-го и 3-го разряда на формовке бетона:

Таблица 6.18

РАСЧЕТ ОБЩЕЙ И ГРУППОВЫХ ДИСПЕРСИЙ

№ п/п	Рабочие 2-го разряда		№ п/п	Рабочие 3-го разряда	
	Дневная выработка (м³) x_1	x_1^2		Дневная выработка (м³) x_2	x_2^2
1	3,2	10,24	1	3,9	15,21
2	3,5	12,25	2	4,2	17,64
3	4,5	20,25	3	4,8	23,04
4	4,8	23,04	4	5,1	26,01
			5	5,4	29,16
			6	6,6	43,56
Итого	16,0	65,78	Итого	30,0	154,62

Отсюда

$$\bar{x}_1 = \frac{16,0}{4} = 4; \quad \bar{x}_2 = \frac{30,0}{6} = 5; \quad \bar{x} = \frac{16 + 30}{10} = 4,6.$$

$$\sigma_1^2 = \bar{x}_1^2 - x_1^2 = \frac{65,78}{4} - 4^2 = 0,445; \quad \sigma_2^2 = \frac{154,62}{6} - 5^2 = 0,77.$$

Общая дисперсия составит:

$$\sigma^2 = \frac{65,78 + 154,62}{10} - (4,6)^2 = 0,88,$$

а средняя из групповых дисперсий равна:

$$\bar{\sigma}^2 = \frac{0,445 \times 4 + 0,77 \times 6}{10} = 0,64.$$

Следовательно, межгрупповая дисперсия должна составить:

$$s^2 = \sigma^2 - \bar{s}^2 = 0,88 - 0,64 = 0,24.$$

Это подтверждается расчетом по формуле 6.27:

$$s^2 = \frac{(4,0 - 4,6)^2 \times 4 + (5,0 - 4,6)^2 \times 6}{10} = 0,24.$$

Правило сложения дисперсий используется в ряде разделов статистики, в частности при расчете ошибок выборочного наблюдения (см. гл. VII) и при измерении тесноты связи между признаками единиц совокупности (см. гл. VIII).

5. АНАЛИЗ ВАРИАЦИОННЫХ РЯДОВ

Понятие о закономерностях распределения

В вариационных рядах, приведенных в гл. IV и в предыдущих параграфах этой главы (см. табл. 4.5, 4.16, 4.17, 6.13 и др.) можно заметить определенную зависимость признака, выражающую закономерность распределения. Например, с увеличением стажа работы численность рабочих и служащих в нефтеперерабатывающей промышленности СССР сначала растет, а затем уменьшается (см. табл. 6.13). Такая же закономерность наблюдается и в отношении распределения колхозов по поголовью коров (см. табл. 4.18).

Совершенно по-иному выглядит распределение семей в СССР по числу совместно проживающих членов (см. табл. 4.17). Здесь по мере возрастания значения вариантов численность семей непрерывно уменьшается. Такой же характер имеет распределение населения по возрасту и ряд других общественных явлений.

Возможно и иное выражение закономерности распределения, что подробно рассматривается в курсе математической статистики. То или иное распределение частот в вариационных рядах определяется двумя обстоятельствами: во-первых, основными причинами, влияющими на характер распределения, и, во-вторых, случайными для данного явления факторами, влияние которых может вызвать более или менее существенные отклонения от закономерного распределения.

Например, распределение рабочих-сдельщиков промышленного предприятия по размеру месячной заработной платы определяется в первую очередь теми причинами, которые заложены в основу сдельной формы оплаты труда. Распределение же рабочих-повременщиков промышленного предприятия по размеру месячной заработной платы определяется в первую очередь теми причинами, которые заложены в основу повременной формы оплаты труда.

При анализе распределения рабочих-сдельщиков промышленных предприятий по проценту выполнения нормы выработки различают две группы с характерными для них особенностями труда и

принципами установления норм выработки: сдельщиков основного производства и сдельщиков вспомогательного производства. Распределение частот в каждой из этих групп имеет свои особенности.

Следовательно, говоря о закономерностях распределения, следует в виду закономерности, присущие качественно однородной совокупности, вариационному ряду, построенному на основе научно обоснованной группировки. В этом выражается связь между методами группировки и закономерностями распределения частот. И так же как недопустимо включение в статистическую совокупность качественно неоднородных единиц, так же лишено экономического смысла определение закономерностей распределения для вариационных рядов, состоящих из качественно неоднородных единиц. Например, закономерности распределения рабочих промышленности (оплачиваемых сдельно и повременно вместе) по размеру месячной заработной платы, так как в этом случае происходит смешение двух разнотипных совокупностей, каждая из которых имеет свою закономерность распределения.

Таким образом, обязательным предварительным условием анализа вариационных рядов является установление качественной однородности изучаемой совокупности в отношении данного признака.

Вместе с тем и в вариационных рядах, составленных из единиц качественно однородной совокупности, закономерность распределения испытывает на себе влияние случайных для данного процесса факторов, что приводит к более или менее существенным отклонениям от закономерного распределения. Поэтому задачей анализа вариационных рядов является установление присущей изучаемому явлению закономерности распределения при элиминировании влияния случайных для данного явления факторов.

Из гл. IV известно, что графическими изображениями вариационных рядов являются **полигон распределения** (см. рис. 4.7) и **гистограмма распределения** (см. рис. 4.8).

Эти графики дают известное представление о закономерности распределения изучаемого явления. Они характеризуют так называемое **эмпирическое распределение**, т. е. то распределение, которое, являясь результатом группировки материалов статистического наблюдения, отражает и основные причины, влияющие на характер распределения и случайные для данного явления факторы.

Обозначение закономерности распределения изучаемого явления и влияния случайных факторов может быть достигнуто путем выявления объема изучаемой совокупности при одновременном увеличении интервалов. В математической статистике доказано, что в результате увеличения объема совокупности и уменьшения промежутков полигона распределения все более и более приближаются к некоторой плавной линии, которая является для него **предельной**. Эта кривая и носит название **кривой распределения**.

Кривая распределения выражает *теоретическое распределение*, т. е. такое распределение, которое в отличие от эмпирического распределения складывается под влиянием только основных, существенных для изучаемого явления причин, при исключении влияния случайных факторов. Это позволяет использовать кривые распределения для изучения законов распределения частот, сравнения распределений и т. д.

Соответственно разнообразным типам закономерностей общественных явлений распределение частот вариационных рядов принимает тот или иной характер, ту или иную форму. В практике статистических исследований широко применяются различные типы кривых распределений: нормальное распределение, распределение Пуассона, распределение Максвелла и др.

Нормальное распределение

Одним из примечательных типов распределения, широко применяемом в экономическом анализе, является нормальное распределение Гаусса — Ляпунова, выражающееся формулой

$$y_t = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}t^2}, \quad (6.29)$$

где y_t — ордината кривой нормального распределения (частости);
 $\pi \approx 3,1415...$
 $e \approx 2,7182...$

t — нормированное отклонение, равное $\frac{x - \bar{x}}{\sigma}$.

Следовательно, кривая нормального распределения может быть построена по двум параметрам: средней арифметической (\bar{x}) и среднему квадратическому отклонению (σ).

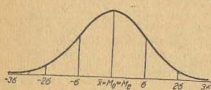


Рис. 6.1 Кривая нормального распределения

Кривая нормального распределения (см. рис. 6.1) представляет одностороннюю симметричную фигуру. От максимального значения, находящегося в середине, кривая равномерно и симметрично убывает в обе стороны, асимптотически приближаясь к оси абсцисс, и образует две равные и подобные ветви. Отсюда

следует, что в нормальном распределении средняя арифметическая, медиана и мода совпадают и находятся в центре (по оси абсцисс).

Как показано на рис. 6.1, кривая нормального распределения имеет две точки перегиба, которые находятся по обе стороны от центра (по оси абсцисс) на расстоянии, равном среднему квадратическому отклонению. В пределах одного среднего квадратического отклонения, т. е. в пределах $\bar{x} \pm \sigma$, при нормальном распределении заключено 68,3% всех значений распределения. В пределах

двух σ , т. е. в пределах $\bar{x} \pm 2\sigma$, заключено 95,4% всех частот и, наконец, в пределах трех σ от центра (по оси абсцисс), т. е. в пределах $\bar{x} \pm 3\sigma$, заключено 99,7% всех частот ряда распределения.

Использование кривой нормального распределения при анализе вариационных рядов связано с тем, что в этом распределении выражается закономерность, складывающаяся под воздействием внутренних причин, вытекающих из сущности данного явления и являющихся для всей совокупности однородных единиц, при взаимоположении причин внешних, индивидуальных, обусловленных случайным для каждой единицы стечением условий и обстоятельств.

Для оценки эмпирического распределения и установления, в какой мере оно выражает закономерности, присущие данному явлению в силу основных причин, составляется кривая нормального распределения, а затем производится сопоставление частот эмпирического ряда с соответствующими частотами кривой нормального распределения.

Возьмем, например, распределение произведенного колхозами урожая по уровню себестоимости 1 т (см. табл. 6.8) и произведем сравнение эмпирических частот по уравнению нормальной кривой (формула 6.29).

Таблица 6.19

ПОСТРОЕНИЕ КРИВОЙ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ

x	m_x	$x - \bar{x}$	$t = \frac{x - \bar{x}}{\sigma}$	$b(t)$	m_t
1	2	3	4	5	6
112,5	9	-11,95	-1,81	0,0775	5,9
117,5	18	-6,95	-1,05	0,2299	17,4
122,5	24	-1,95	-0,30	0,3814	28,9
127,5	27	3,05	0,46	0,3589	27,2
132,5	18	8,05	1,22	0,1895	14,4
137,5	4	13,05	1,98	0,0562	4,3
Итого	100	—	—	—	98,1

В гр. 1 приведены центральные эмпирические соответствующих интервалов из табл. 6.8, в гр. 2 — эмпирические частоты этих интервалов. В гр. 3 показаны разности между центральными эмпирическими вариантами и средним арифметическим значением признака, равным 124 р. 45 к. (см. с. 167). В гр. 4 приведены значения нормированных отклонений (t), исчисленных путем деления разностей между центральными эмпирическими вариантами и средним арифметическим значением признака на среднее квадратическое отклонение, равное 6 р. 59 к. (см. с. 181). Величина (b) определяется по таблицам, имеющимся в учебниках математической статистики. Наконец, теоретические значения кривой

нормального распределения, приведенные в гр. 6, исчислены по формуле

$$m_i = \frac{f(t) \cdot h \cdot \sum m}{\sigma}, \quad (6.30)$$

где h — ширина интервала и $\sum m$ — объем совокупности. Например,

$$m_{T_1} = \frac{0,0755 \times 5 \times 100}{6,59} = 5,9;$$

$$m_{T_2} = \frac{0,2299 \times 5 \times 100}{6,59} = 17,4$$

и т. д.

Теоретическая сумма частот оказалась равной 98,1, а не 100 в результате округлений в расчетах.

Как видно из таблицы, эмпирические (фактические) частоты и теоретические (исчисленные) частоты достаточно близки, хотя в отдельных случаях между ними имеются расхождения. Более наглядное сопоставление достигается с помощью соответствующего графика: две кривые на одном чертеже (см. рис. 6.2).

Критерии соответствия или согласия

уравнению нормальной кривой, имеются несовпадения. В чем их причина?

Несовпадения между фактическими и теоретическими частотами могут иметь две причины. Во-первых, некоторое расхождение фактических и теоретических частот может быть чисто случайным, оно может быть вызвано факторами случайными для данного явления. Во-вторых, расхождение между фактическими и теоретическими частотами может быть существенным, его причиной может быть несоответствие изучаемого распределения нормальному по своей природе.

Чтобы можно было объективно судить, какая именно причина привела к расхождению между фактическими и теоретическими частотами, необходим некоторый критерий соответствия, критерий согласия. Это должен быть такой пока-

затель, который дал бы возможность, опираясь на закон распределения определенного вида, установить, когда отклонения фактических частот от теоретических следует считать существенными и когда несущественными, случайными. В статистических исследованиях применяются критерии согласия Пирсона, Романовского, Колмогорова и Ястремского, в основе которых лежат те или иные теоретические положения.

Так, критерий согласия, разработанный советским ученым академиком А. Н. Колмогоровым, **критерий лямбда** (λ) основан на сопоставлении сумм накопленных эмпирических (фактических) и теоретических (исчисленных) частот и исчисляется по формуле

$$\lambda = \frac{D}{\sqrt{\sum m}}, \quad (6.31)$$

где D — максимальное значение разности между накопленными эмпирическими и теоретическими частотами;

$\sum m$ — сумма эмпирических частот.

И в качестве примера рассчитаем критерий согласия лямбда по данным табл. 6.19 об эмпирических и теоретических частотах при допущении, что сумма эмпирических частот в абсолютном выражении равна 100.

Таблица 6.20

РАСЧЕТ КРИТЕРИЯ СОГЛАСИЯ ЛАМБДА

Интервалы классов (абсолютные частоты) f (гр. 6)	Частоты ряда распределения		Накопленные частоты		Абсолютное значение разности накопленных эмпирических и теоретических частот $ \sum m_2 - \sum m_1 $
	эмпирические m_2	теоретические m_1	эмпирические $\sum m_2$	теоретические $\sum m_1$	
110—115	9	5,9	9	5,9	3,1
115—120	18	17,4	27	23,3	3,7
120—125	24	28,9	51	52,2	1,2
125—130	27	27,2	78	79,4	1,4
130—135	18	14,4	96	93,8	2,2
135—140	4	4,3	100	98,1	1,9
Итого	100	98,1*	×	×	×

* Теоретическая сумма частот оказалась равной 98,1, а не 100 в результате округлений в расчетах.

Из таблицы видно, что максимальное значение разности между эмпирическими и теоретическими частотами равно 3,7. Отсюда величина критерия согласия лямбда составляет:

$$\lambda = \frac{3,7}{\sqrt{100}} = \frac{3,7}{10} = 0,37.$$

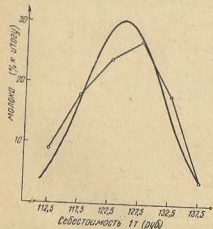


Рис. 6.2. Распределение произведенного молока по уровню себестоимости 1 т

С помощью специально разработанных таблиц определена вероятность наличия близости между эмпирическим и теоретическим распределениями при том или ином значении λ (см. табл. 6.21). Как видно из табл. 6.21, значения вероятностей λ для колеблются в пределах от 0 до 1. При полном совпадении эмпирического и теоретического распределений вероятность λ равна 1 и, наоборот, при полном расхождении между эмпирическим и теоретическим распределениями вероятность λ равна 0.

ТАБЛИЦА ВЕРОЯТНОСТЕЙ ДЛЯ КРИТЕРИЯ СОГЛАСИЯ ЛАМБДА

λ	$P(\lambda)$	λ	$P(\lambda)$	λ	$P(\lambda)$
0,30	1,000	1,00	0,270	1,70	0,006
0,40	0,997	1,10	0,178	1,80	0,003
0,50	0,964	1,20	0,112	1,90	0,002
0,60	0,864	1,30	0,068	2,00	0,0007
0,70	0,711	1,40	0,040	2,10	0,0001
0,80	0,544	1,50	0,022	2,20	0,0001
0,90	0,393	1,60	0,012	2,31	0,0000

Таблица 6.21

Так, для рассматриваемого примера вероятность критерия согласия λ очень близка к 1, что означает, что с вероятностью, близкой к полной достоверности, можно считать отклонения эмпирических частот от теоретических незначительными. Рассматриваемое распределение полностью согласуется с нормальным распределением.

Для анализа вариационных рядов, имеющих иной, сравнительно с нормальной кривой, характер распределения, математической статистикой разработаны соответствующие формулы расчета теоретических значений частот.

Асимметрия распределения и эксцесс

Критерии согласия, дающие общую оценку степени совпадения эмпирических частот с вычисленными по уравнению нормальной кривой, не показывают, однако, чем конкретно отличается эмпирическое распределение от нормального. Для этой цели применяются специально разработанные показатели *асимметрии* и *эксцесса*.

Ранее было отмечено, что кривая нормального распределения представляет одновершинную симметричную фигуру. Сопоставляя эмпирическое распределение с нормальным, прежде всего устанавливали, в какой мере оно отклоняется по симметричности, либо, что то же, в какой мере оно *асимметрично*.

Если в симметричном распределении вершина кривой находится в центре, то при асимметрии вершина кривой сдвинута от центра либо вправо — левосторонняя (отрицательная) асимметрия, либо влево — правосторонняя (положительная) асимметрия. Если

в симметричном распределении частоты любых двух вариантов, состоящих в обе стороны от среднего значения признака, равны между собой, то в асимметричном распределении эти частоты с одной стороны все время более или все время менее соответственных частот с другой стороны. Наконец, если в симметричном распределении средняя арифметическая, медиана и мода совпадают, то в асимметричном распределении они не совпадают. При левосторонней (отрицательной) асимметрии $Mo > Me > x$, а при правосторонней (положительной) асимметрии $Mo < Me < x$.

Степень асимметрии измеряется с помощью следующего коэффициента:

$$K_A = \frac{\bar{x} - Mo}{\sigma}, \quad (6.32)$$

где коэффициент асимметрии представляет частное от деления разности между средней арифметической и модой на среднее квадратическое отклонение.

При левосторонней (отрицательной) асимметрии коэффициент асимметрии имеет знак — (минус), при правосторонней (положительной) асимметрии этот коэффициент имеет знак + (плюс).

Так, применительно к ранее рассмотренному примеру распределения произведенного колхозами молока по уровню себестоимости

$$K_A = \frac{124,45 - 126,25}{6,59} = -0,27,$$

это говорит о незначительной по величине и отрицательной по знаку характеру асимметрии.

Следующим шагом анализа эмпирического распределения является определение его *эксцесса*, т. е. островершинности сравнительно с кривой нормального распределения. Математической статистикой установлено, что для кривой нормального распределения отношение четвертого центрального момента (μ_4) к четвертой степени среднего квадратического отклонения (σ^4) равно 3. Отсюда коэффициент эксцесса представляет разность

$$E = \frac{\mu_4}{\sigma^4} - 3. \quad (6.33)$$

Отношение $r_4 = \mu_4 : \sigma^4$ называется мерой крутости распределения.

Для кривой нормального распределения $r_4 = 3$ и $E = 0$.

Показатель эксцесса может быть положительным, что характеризует островершинность кривой распределения по сравнению с нормальной кривой, и отрицательным, что характеризует плосковершинность кривой распределения по сравнению с нормальной кривой.

6. КРИТИКА БУРЖУАЗНЫХ «ТЕОРИЙ» СРЕДНИХ И ПРИЕМОВ ИХ ВЫЧИСЛЕНИЯ

Трактовка средней величины в буржуазной статистике Основанная на идеалистической философии, имеющая ярко выраженный формально-математический характер, буржуазная статистика не в состоянии дать правильное объяснение сущности и выбора формы средних величин, а это создает почву для ненаучных приемов их вычисления.

В буржуазной статистике существуют разные точки зрения на природу и сущность средних величин. Согласно одной из них любая средняя величина является отражением некоторой «истинной» величины, недоступной непосредственному измерению.

Сторонники этой теории, которая в литературе известна как теория истинных величин, исходят из того, что в ряде случаев при многократном измерении одной и той же величины, например расстояния между двумя точками местности, результаты измерений могут быть различными, и в этих случаях наиболее точно размер искомой величины выражает среднее значение отдельных измерений. Действительно, в средней, полученной из большого числа наблюдений, взаимно погашаются ошибки наблюдения, возникающие как вследствие неточности инструментов, так и в результате ошибок наблюдателя. Здесь находит проявление закон больших чисел. Но из этих правильных для частного случая посылок делается логически не обоснованный вывод, будто любая средняя величина является отражением некоторой «истинной» величины, а индивидуальные величины, из которых исчисляется средняя, следует рассматривать как случайные отклонения от этой «истинной» величины.

С позиций идеалистической философии авторы этой теории утверждают, что подобно тому, как действительный мир есть будто бы лишь отражение мира идей, так и существующий в нашем мире конкретные величины есть лишь отображение идеально существующих «истинных» величин. Мир реально существующих величин, т. е., по существу, все явления конкретного мира, нужно, по мысли сторонников этой теории, рассматривать как случайные отклонения от «истинных» величин, содержащие ту или иную ошибку. А так как отклонения в каждом отдельном случае одинаково вероятны как в одну, так и в другую сторону, то при достаточно большом числе наблюдений эти отклонения взаимно погасятся в средней. Найденная таким путем средняя или совпадает с «истинной» величиной, или будет незначительно от нее отличаться. Отсюда делается вывод, что средняя величина — некоторое более или менее точное приближение к идеально существующим «истинным» величинам.

Основной порок теории «истинных» величин заключается в ненаучном, идеалистическом взгляде на природу вещей. Именно это предопределяет игнорирование качественного анализа изучаемых общественных явлений при исчислении средних. Придание же

средним характера абсолютной неизменности, устойчивости делает эту теорию реакционной, ставит ее на службу капитализма.

Основателем теории истинных величин и ее видимым представителем является бельгийский астроном, математик и статистик XIX в. Адольф Кетле (1796—1874), который известен также как автор теории «среднего человека». Кетле писал: «Рассматриваемый абстрактно, как представитель всего нашего рода и как носитель наших качеств, какие только встречаются у других людей, человек будет называться у нас средним человеком»¹.

Средний человек у Кетле имеет средний рост и вес, среднюю склонность к браку, самоубийству, к преступлениям и добрым делам и т. п. По мысли Кетле, каждый конкретный человек является как бы случайным отклонением от «среднего человека», который всеми качествами наделен в среднем размере.

Механически отождествляя законы природы с законами общественного развития, Кетле утверждает далее, что «можно заранее сказать, сколько индивидуумов замарают руки в крови своих ближних, сколько явится делателей фальшивых бумаг, сколько отравителей и пр.»².

Легко видеть, что эта идеалистическая трактовка средних величин применительно к человеку ненаучна и реакционна. Вполне закономерно, что она вызвала резкую критику со стороны прогрессивных ученых XIX в. В частности, видный русский статистик Ю. Э. Янсон, вскрывая несостоятельность теоретического обоснования, на котором Кетле построил свою теорию, отмечал, что «предположение Кетле о существовании в природе среднего человека как чего-то типичного, от которого жизнь отклоняла реально существующих людей, приводит к механическому взгляду на законы движения социальной жизни и отрицанию поступательного движения общества».

Теория распространенная среди буржуазных статистиков трактовка средних величин основана на философии махизма (одна из разновидностей идеалистической философии) с характерным для нее принципом «экономии мышления». Например, основатель этой теории английский профессор статистики А. Боули пишет: «Ум не в состоянии сразу охватить величины миллионов статистических данных, они должны быть сгруппированы, упрощены, приведены к средним»³. Следовательно, замена множества величин одной средней в целях «экономии мышления» — такова будто бы задача статистики.

Взгляд на метод средних как на технический прием упрощения массовой информации широко распространен и среди других представителей англо-американской школы статистиков (Р. Фишер, Л. Миддлс, Дж. Юл и др.). Например, Р. Фишер полагает, что «такая человеческий ум не способен вместить в себя все содер-

¹ Кетле А. Социальная физика, т. 1, Киев, 1911, с. 52.

² Кетле А. Человек, развитие его способностей или опыт социальной физики, СПб., 1865, с. 7.

³ Боули А. Элементы статистики, ч. 1, 1937, с. 125.

жаке более или менее значительного количества числовых данных. Поэтому мы всегда стремимся к тому, чтобы отразить в относительно небольшом числе сводных показателей наиболее важную и существенную информацию, содержащуюся в данной массе наблюдений. Все это является простой практической необходимостью, и теория статистики с этим должна считаться¹.

И эта разновидность буржуазной трактовки средних является ненаучной, так как средняя величина у сторонников этой теории выступает не как обобщающая характеристика однокачественной совокупности, а как чисто счетная категория, не имеющая никакого объективного значения и познавательной ценности. Отсюда формально-математическая трактовка средних, широкое применение огульных средних, являющихся удобной основой для различного рода фальсификаций в целях апологетики капитализма.

Фальсификация средних в буржуазной статистике

В соответствии с ненаучными, идеалистическими теориями средних буржуазной статистикой применяются фальсификаторские приемы их исчисления. Игнорируя требование качественной однородности совокупностей, для которых исчисляются средние, буржуазная статистика широко применяет огульные средние как средство обмана обществу и в первую очередь широких трудящихся масс. Так исчисляется средняя продолжительность жизни для всего населения страны, чтобы скрыть, что в условиях капитализма особенно высокая смертность среди трудящихся, а смертность так называемого высшего населения во много раз превышает смертность белого населения.

Буржуазная статистика обычно хвастает высокими средними доходами, смазывая тот факт, что ничтожная кучка богатеев получает миллионные прибыли, а подавляющей массе населения не хватает средств на существование.

Например, официальные цифры средней заработной платы рабочих и служащих увеличиваются путем включения в общую сумму заработной платы собственно рабочих и служащих оклады директоров акционерных компаний и прочих высокооплачиваемых лиц. Так, по данным американской печати, председатель корпорации «Дженерал Моторс» Ф. Доннер получает в год оклад и надбавки в сумме 574 тыс. долларов; Ш. Скинер, вице-президент «Дженерал Моторс», — 517 тыс. долларов и т. д. Конечно, если подобного рода «заработную плату» присоединять к заработной плате рабочих и служащих, то можно получить достаточно высокую «среднюю» заработную плату. Но чего стоят такая «средняя»?

С помощью подобной нехитрой арифметики американская буржуазная статистика создала миф о «среднем американце» и его

существовании. Но нетрудно понять, что в действительности никакого «среднего американца» не существует, а есть американский капиталист и американский калиталист.

Задачей советской статистики является разоблачение фальсификаторских ненаучных приемов исчисления средних, применение буржуазной статистики.

Выбор формы средней величины

Формально-математическая трактовка средних (в отрыве от качественного анализа) скрывается и на решении буржуазной статистики вопроса о выборе формы средней.

Например, Ф. Миллс пишет, что средняя арифметическая должна применяться, «если абсолютные цифры, нанесенные на арифметическую шкалу, дают совершенно симметричное распределение». Но когда нанесенные на диаграмму абсолютные цифры дают асимметричную кривую и при этом такого типа, что если вместо самих чисел их логарифмы, то асимметрия кривой исчезает, тогда более предпочтительна средняя геометрическая². Что же касается средней гармонической, то это особый вид средней, применимый только в ограниченной области; ею можно пользоваться при выводе средних норм времени и она дает заметные выгоды при обработке некоторых данных о цене³.

Р. Дюма считает, что «применение средней гармонической объяснимо в тех случаях ..., когда показатели не могут быть сложены». А в учебнике Дж. Э. Юла и М. Дж. Кендэла «Теория статистики», выдержавшем многократные издания, о гармонической средней говорится лишь, что «гармоническая средняя ряда имеет ту же величину, обратная арифметической средней из обратных чисел»⁴. Приводимый затем пример может служить не столько иллюстрацией способа вычисления средней гармонической, сколько иллюстрацией формализма в статистике.

Только подчинив выбор формы средней социально-экономическому содержанию изучаемых явлений, можно найти однокачественное значение средней как типической характеристики признаков изучаемых явлений.

¹ Миллс Ф. Статистические методы (Пер. с англ.). М., Госстатиздат, 1958, с. 107—108.

² Там же, с. 110.

³ Дюма Р. Предприятие и статистика (Пер. с франц.). М., Госстатиздат, 1958, с. 10.

⁴ Юл Дж. Э. и Кендэл М. Дж. Теория статистики (Пер. с англ.). М., Госстатиздат, 1960, с. 155.

¹ Фишер Р. А. Статистические методы для исследователей (Пер. с англ.). М., Госстатиздат, 1958, с. 14.

ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

Примени-
мость выбо-
рочного
наблюдения

Имеется ряд причин, в силу которых во многих случаях выборочному наблюдению отдается предпочтение перед сплошным. Из них наиболее существенны следующие:

1. *Экономия времени и средств* в результате сокращения объема работы. Действительно, подвергнув обследованию 1—2% общего числа единиц совокупности, сокращают объем работы в первом варианте примерно в 100 раз, а во втором — в 50. Этим достигается сокращение времени исследования и одновременно уменьшение его стоимости.

При громадном объеме статистических работ в нашей стране уменьшение их стоимости и ускорение сроков их проведения имеют большое народнохозяйственное значение. Вот почему выборочный метод получил такое широкое распространение в статистических исследованиях.

2. *Сведение к минимуму порчи или даже уничтожения исследуемых объектов.* Известно, что проверка качества продукции часто связана с ее порчей либо даже уничтожением. Примерами могут служить испытание пружин на разрыв, испытание электроламп на длительность горения, определение степени созревания сахарной свеклы путем ее выборочной копки либо всхожести семян той или иной культуры и т. д. Естественно, что при подобном виде исследований применяется только выборочное наблюдение, так как сплошное наблюдение привело бы к бессмысленному уничтожению плодов человеческой деятельности.

3. *Необходимость детального исследования каждой единицы совокупности при невозможности охвата всех единиц.* Например, для изучения бюджетов семей трудящихся требуется, чтобы в отдельных семьях велся точный и систематический учет доходов и расходов, чтобы эти записи велись ежедневно. Совершенно ясно, что такая работа не может быть организована во всех семьях страны, а должна производиться лишь в относительно небольшом числе семей.

4. *Достижение большей точности результатов обследования благодаря сокращению ошибок, происходящих при регистрации.* Ошибки регистрации (см. гл. III) возможны как при сплошном, так и при выборочном наблюдении. Но установлено, что разность этих ошибок, как правило, больше при сплошном наблюдении. Это объясняется тем, что при выборочном наблюдении благодаря сокращению количества обследуемых единиц возможна более квалифицированная и тщательная регистрация.

В то же время при выборочном наблюдении возникают ошибки, которые не имеют места при сплошном наблюдении, — *ошибки репрезентативности.* Эти ошибки — результат того, что состав выборочной совокупности обычно в какой-то мере отличается от состава генеральной совокупности, а это приводит к тому, что обобщающие выборочные показатели отличаются от соответствующих генеральных показателей. Разность между обобщающими

1. ОСНОВНЫЕ ПОЛОЖЕНИЯ ВЫБОРОЧНОГО МЕТОДА

Общее понятие
о выборочном
наблюдении

Выборочное наблюдение является одним из видов несплошного наблюдения. При выборочном наблюдении исследованию подвергается некоторая часть совокупности, а обобщающие показатели, характеризующие эту исследованную часть совокупности, распространяются на всю совокупность.

Последний момент является очень важным для понимания сущности выборочного наблюдения. Если некоторая часть единиц совокупности изучается для того, чтобы вывести обобщающие показатели, характеризующие только эту часть совокупности, то такое исследование не будет выборочным. Например, если для изучения вопроса об уровне специализации и кооперирования в промышленности были обследованы не все промышленные предприятия, а только предприятия машиностроения, то полученные в результате этого статистические характеристики уровня специализации и кооперирования можно отнести только к машиностроению. Такое исследование может быть названо частичным, но не выборочным. Если же для изучения специализации и кооперирования были подвергнуты наблюдению, допустим, 10% общего числа промышленных предприятий всех отраслей промышленности, то в этом случае были бы получены такие статистические характеристики, которые можно распространять на всю промышленность. Такое исследование является выборочным.

Общая совокупность единиц, из которой производится отбор, называется *генеральной*. Отобранная определенным образом часть генеральной совокупности, подлежащая выборочному обследованию, называется *выборочной* совокупностью. Обобщающие показатели генеральной совокупности (средняя, доля, дисперсия) называются *генеральными*, а соответствующие обобщающие показатели выборочной совокупности — *выборочными*.

Выборочный метод отличается от других видов несплошного наблюдения (мониторинга, описаний, анкетного метода и метода основного массива, особенности которых изложены в гл. III) двумя признаками: 1) заранее устанавливается, сколько единиц или какая часть единиц генеральной совокупности будет обследована, и 2) заранее определяется порядок отбора единиц, при котором выборочная совокупность в достаточной мере представляла бы (репрезентовала) генеральную совокупность.

выборочными показателями и соответствующими показателями генеральной совокупности представляет собой ошибку репрезентативности.

Конечно, если бы выборочная совокупность точно представляла генеральную, то частоты в ней, отличаясь от частот генеральной совокупности в буквальном выражении, совпали бы по абсолютным весам в общем итоге. В этом случае выборочные обобщающие показатели совпали бы с генеральными. Однако вероятность этого очень мала. Как правило, выборочная совокупность не в полной мере воспроизводит состав генеральной совокупности. Поэтому выборочные обобщающие показатели (средняя, доля, дисперсия) обычно отличаются от соответствующих показателей генеральной совокупности.

Ошибки репрезентативности делятся на *случайные* и *систематические*. Случайные ошибки не имеют преимущественного направления в сторону преувеличения или преуменьшения величин изучаемого показателя. Систематические же ошибки направлены в одну определенную сторону. Например, если при выборочной копке сахарной свеклы для определения ее урожайности в выборку случайно, в силу ряда причин, попадут несколько лучших, чем в среднем, экземпляров, то речь будет идти о случайной ошибке репрезентативности. Если же при выборке не случайно, а с какими-нибудь предвзятыми целями будут систематически отбираться только лучшие или только худшие экземпляры, то здесь речь уже будет идти о систематической ошибке репрезентативности, получившейся в результате преднамеренного нарушения правил отбора. Для предупреждения и устранения ошибок репрезентативности нужно установить научно обоснованный порядок отбора, исключающий как преднамеренное, так и непреднамеренное искажение размеров изучаемых признаков.

При соответствующей организации выборки ошибки репрезентативности становятся настолько незначительными, а ошибки регистрации настолько сокращаются по сравнению со сплошным наблюдением, что выборочное наблюдение может стать более точным, чем иное сплошное. Следовательно, глубоко ошибочным является мнение, будто всякое сплошное наблюдение точнее выборочного. Как подтверждает практика статистических работ, выборочный метод иногда применяется для контроля точности результатов сплошного наблюдения (например, выборочные контрольные обходы по проверке численности скота, находящегося в личной собственности населения, проживающего в сельской местности).

Закон больших чисел и выборочный метод

Ответ на вопрос о том, как велика и насколько вероятна разница между генеральными и выборочными обобщающими показателями, дает теория выборочного метода, основанная на законе больших чисел. Этот ответ она дает на основе решения двух взаимосвязанных задач: во-первых, расчета с заданной вероятностью предела возможных отклонений выборочного показателя от соответствующего показателя в генеральной совокупности и, во-вторых, определения ва-

риатности того, что размер возможных отклонений выборочного показателя от генерального не превысит заданного предела.

Большое значение для развития теоретических основ выборочного метода имела работа русских математиков в области теории вероятностей. Со второй половины XIX в. в разработке научных основ выборочного метода связаны с именами выдающихся ученых-математиков академиков П. Л. Чебышева, А. М. Ляпунова, А. А. Маркова, а в советский период — с именами крупнейших советских математиков академиков А. Н. Колмогорова, С. Н. Бернштейна, а также А. Я. Хинчина, В. В. Гнеденко, В. И. Романовского и др.

Большую роль в теоретическом обосновании выборочных обследований обобщающих явлений имели работы советских статистиков Э. С. Немычина, В. В. Петровского, В. Н. Старовского, А. Я. Боярского и др.

СПОСОБЫ ОТБОРА.

ОТБОР ОБЕСПЕЧИВАЮЩИЕ РЕПРЕЗЕНТАТИВНОСТЬ ВЫБОРКИ

Возможны различные способы отбора единиц исследуемой генеральной совокупности в целях образования выборочной совокупности. Прежде всего нужно отличать *индивидуальный* и *серийный отборы*. При индивидуальном отборе выборочная совокупность образуется путем последовательного отбора отдельных единиц при *серийном* — путем последовательного отбора целых единиц (*серий*) единиц, после чего обследуются все единицы отобранных групп. В свою очередь индивидуальный и серийный отборы могут быть организованы как *собственно-случайный*, *механический* и *типический отбор*. Часто эти способы отбора сочетаются с целью достижения более высокой репрезентативности. Так, *серийный отбор* может сочетаться с *типическим*, *типический* — с *механическим* и т. д.

Собственно-случайный, *типический* и *серийный* способы отбора могут быть организованы в виде *повторной* и *бесповторной* выборки. При повторном отборе каждая единица (или группа единиц) генеральной совокупности, попавшая в выборку, после записи размера интересующего нас признака вновь возвращается в генеральную совокупность и, следовательно, может многократно попасть в выборку. При бесповторном же отборе каждая единица (или группа единиц) генеральной совокупности, попавшая в выборку, после записи размера интересующего нас признака исключается из генеральной совокупности не возвращается. Выбор единичного или многократного способа отбора зависит от характера изучаемых общественных явлений и стоящей перед исследователем задачи.

Бесповторный отбор применяется чаще, например, при выборочном изучении качества продукции, размера потерь при уборке урожая и т. д., где повторный отбор практически нецелесообразен. Повторный же применяется при изучении потребительского спроса в торговле, производится обследование пассажиров трамвая и пригородных поездов и т. д., так как бесповторный отбор здесь практически неосуществим.

Таблица 7.1

СВОБОДНЫЙ СТАЖ И КВАЛИФИКАЦИЯ РАБОЧИХ СБОРОЧНОГО ЦЕХА

Собственно-случайная выборка

Собственно-случайной называется такая выборка, при которой отбор единиц (или групп единиц) для обследования производится из всей генеральной совокупности непреднамеренно и случайным порядком. Очень часто с этой целью применяется жребьевка.

Например, в сборочном цехе машиностроительного завода работают 300 рабочих, объединенных в 30 бригад, по 10 человек в каждой бригаде (см. табл. 7.1; сведения о рядах и порядковых номерах рабочих при группировке по рядам будут использоваться при рассмотрении типического отбора).

Необходимо определить средний стаж рабочих и долю (процент) рабочих со стажем более 10 лет.

Общая сумма проработанных всеми рабочими человеко-лет равна 2700, откуда средний стаж составляет $x = 2700 : 300 = 9$ (лет).

Рабочих со стажем свыше 10 лет 93 человека, откуда доля их в общей численности рабочих $p = 93 : 300 = 0,31$, или 31%.

Теперь посмотрим, какие бы мы получили результаты, если бы, не имея полных данных, взяли за решение поставленной задачи на основе десятипроцентной собственно-случайной повторной выборки. Для этого необходимо было бы проделать следующее:

- 1) составить список рабочих цеха по порядку табельных номеров (от 1-го до 300-го номера) с проставлением стажа;
- 2) изготовить из бумаги одинакового размера билеты (жребии), подписать их номерами от 1 до 300, проставить на них соответствующий стаж рабочего, свернуть билеты в трубочки и вложить в специальные патроны;
- 3) поместить патроны с билетами в ящик (урну) и тщательно перемешать их;
- 4) вынуть из урны один билет, записать проставленный на нем стаж, после чего положить его снова в урну, опять перемешать билеты, после чего вынуть второй билет и т. д., а всего вынуть таким образом 30 билетов (10% от 300).

Рабочие, чьи табельные номера будут значиться на вынутых билетах, составят выборочную совокупность. Как видно из этого, при случайном отборе для каждой единицы создаются условия равной возможности попасть в выборку в каждом отдельном акте отбора. Благодаря этому при достаточно большом объеме выборки обобщающие показатели выборочной совокупности должны довольно точно воспроизвести обобщающие показатели генеральной совокупности.

Допустим, что в выборку попали рабочие, имеющие следующие табельные номера: 116, 54, 40, 177, 12, 37, 145, 42, 127, 84, 205, 271, 226, 256, 297, 147, 93, 103, 85, 154, 12, 200, 81, 186, 1, 103, 275, 300, 230, 222.

Стаж рабочих, попавших в выборку (лет): 5, 7, 4, 9, 11, 1, 18, 18, 3, 10, 6, 22, 13, 23, 3, 2, 2, 9, 11, 4, 20, 8, 30, 2, 18, 8, 5, 3.

Ряд	М. п.	Табельный номер	Стаж работы	Ряд	М. п.	Табельный номер	Стаж работы	Ряд	М. п.	Табельный номер	Стаж работы	Ряд	М. п.
1	2	3	4	1	2	3	4	1	2	3	4	1	2
30	6	241	41	28	6	251	81	22	5	250			
13	4	151	42	18	5	252	82	1	1	37			
6	2	1	43	2	2	20	83	9	2	38			
4	2	2	44	1	1	21	84	10	3	176			
11	3	152	45	3	3	162	85	2	1	39			
10	6	242	46	17	4	163	86	14	4	40			
11	3	153	47	1	1	22	87	14	4	177			
1	1	1	3	48	10	3	164	88	1	41			
12	4	154	49	13	4	165	89	19	5	251			
1	5	243	50	1	2	23	90	10	2	42			
3	1	4	51	2	2	24	91	9	2	43			
11	3	155	52	2	1	25	92	1	1	44			
9	2	5	53	18	5	253	93	3	2	45			
21	5	244	54	7	3	166	94	1	1	46			
21	5	245	55	19	4	167	95	4	3	178			
2	1	6	56	6	2	26	96	1	1	47			
10	3	156	57	28	5	254	97	19	6	262			
19	4	157	58	6	2	27	98	3	4	179			
6	2	7	59	3	2	28	99	2	1	48			
18	5	246	60	11	6	255	100	6	5	263			
1	1	8	61	1	1	29	101	3	2	49			
7	2	9	62	6	3	168	102	32	6	264			
21	5	247	63	20	4	169	103	2	1	50			
2	1	10	64	5	2	30	104	1	1	51			
14	3	158	65	33	6	256	105	16	5	265			
9	3	159	66	10	3	170	106	1	1	52			
3	1	11	67	2	1	31	107	2	1	53			
4	2	12	68	2	2	32	108	3	2	54			
19	5	248	69	6	3	171	109	7	3	180			
1	1	13	70	1	1	33	110	12	4	181			
10	3	160	71	7	3	172	111	30	6	266			
8	2	14	72	11	4	173	112	5	4	182			
17	6	249	73	28	6	257	113	13	3	183			
4	2	15	74	19	5	258	114	1	1	55			
1	1	16	75	17	4	174	115	4	2	56			
23	5	250	76	10	3	175	116	5	1	57			
1	1	17	77	7	2	34	117	4	3	184			
1	2	18	78	14	6	259	118	7	2	58			
6	3	161	79	9	2	35	119	6	4	185			
4	1	19	80	8	2	36	120	8	2	59			

А	1	2	3	А	1	2	3	А	1	2	3	А	1	2	3	А	1	2	3
121	15	4	186	161	10	2	81	201	13	4	101	15	6	287	261	3	2	124	281
122	24	5	267	162	3	1	82	202	5	2	102	5	2	116	262	1	1	125	282
123	2	1	60	163	32	5	272	203	12	4	103	8	3	230	263	7	2	126	283
124	7	2	61	164	9	6	273	204	1	1	104	9	4	231	264	10	4	127	284
125	1	1	62	165	19	4	200	205	6	3	105	10	5	232	265	4	2	128	285
126	3	2	63	166	1	83	206	18	4	1	106	12	6	288	266	17	6	129	286
127	3	1	64	167	4	3	201	207	4	2	107	5	2	117	268	5	2	130	288
128	5	3	187	168	18	4	202	208	23	6	108	3	1	118	269	3	1	131	289
129	20	4	188	169	7	2	84	209	2	1	109	9	3	233	270	4	2	132	290
130	2	1	65	170	2	1	85	210	2	1	110								
131	4	2	66	171	3	2	86	211	1	1	111	4	2	119	271	22	5	294	291
132	5	2	67	172	19	4	203	212	1	1	112	5	2	120	272	23	5	295	292
133	4	3	189	173	11	5	274	213	8	3	113	6	3	234	273	8	2	131	293
134	3	1	68	174	2	1	87	214	7	3	114	7	2	235	274	7	2	132	294
135	1	1	69	175	30	6	275	215	9	4	115	8	3	236	275	18	6	296	295
136	21	5	208	176	7	3	204	216	16	4	116	10	4	237	276	4	1	133	296
137	19	4	190	177	9	4	205	217	3	1	117	11	5	238	277	11	3	238	297
138	4	1	70	178	4	2	88	218	22	5	118	12	6	239	278	6	2	134	298
139	4	3	191	179	5	2	89	219	5	2	119	13	7	240	279	6	2	135	299
140	10	4	192	180	1	1	90	220	6	3	120	14	8	241	280	3	1	136	300
141	5	2	71	181	7	2	91	221	2	1	121	15	9	242	281	4	2	137	301
142	6	3	193	182	11	5	276	222	3	2	122	16	10	243	282	5	3	138	302
143	7	2	72	183	21	5	277	223	8	3	123	17	11	244	283	6	4	139	303
144	1	1	73	184	3	1	92	224	15	4	124	18	12	245	284	7	5	140	304
145	18	4	194	185	6	2	93	225	24	6	125	19	13	246	285	8	6	141	305
146	9	2	74	186	8	4	206	226	13	4	126	20	14	247	286	9	7	142	306
147	23	6	269	187	3	1	94	227	5	2	127	21	15	248	287	10	8	143	307
148	12	5	270	188	10	5	278	228	1	1	128	22	16	249	288	11	9	144	308
149	1	1	75	189	1	1	95	229	8	3	129	23	17	250	289	12	10	145	309
150	5	4	195	190	6	2	96	230	5	2	130	24	18	251	290	13	11	146	310
151	7	2	76	191	5	3	207	231	5	2	131	25	19	252	291	14	12	147	311
152	9	3	196	192	1	1	97	232	6	3	132	26	20	253	292	15	13	148	312
153	14	6	271	193	11	5	279	233	15	4	133	27	21	254	293	16	14	149	313
154	9	3	197	194	21	6	280	234	7	3	134	28	22	255	294	17	15	150	314
155	6	2	77	195	18	4	208	235	19	5	281	29	23	256	295	18	16	151	315
156	2	1	78	196	2	1	98	236	21	5	282	30	24	257	296	19	17	152	316
157	2	1	79	197	4	2	99	237	6	3	283	31	25	258	297	20	18	153	317
158	5	3	198	198	8	3	209	238	8	3	284	32	26	259	298	21	19	154	318
159	8	2	80	199	14	4	210	239	10	4	285	33	27	260	299	22	20	155	319
160	10	4	199	200	4	1	100	240	14	5	286	34	28	261	300	23	21	156	320

Сумма этих показателей составляет 300 лет, откуда средняя стаж рабочих в выборочной совокупности $x = 300 : 30 = 10$ (лет).

Так как в выборку попало 11 рабочих со стажем свыше 10 лет, то выборочная доля составляет $w = 11 : 30 = 0,367$, или 36,7%.

Как видим, выборочная средняя отличается от генеральной средней, а выборочная доля — от генеральной доли.

При собственно-случайном отборе выборочная средняя и выборочная доля являются переменными величинами. Они могут принимать различные значения при том или ином исходе выборки — той или иной вероятностью, колеблясь соответственно около значений генеральной средней или доли. Мерой колеблемости возможных значений выборочной средней около генеральной средней, а выборочной доли около генеральной доли является дисперсия, т. е. средний квадрат отклонений. Обозначив эту величину через μ^2 (греческая буква «мю» в квадрате), имеем:

$$\mu_x^2 = \frac{\sum (\bar{x}_i - \bar{x})^2}{n}$$

$$\mu_w^2 = \frac{\sum (w_i - w)^2}{n}$$

В математической статистике доказывается, что для случайной простой выборки между дисперсией выборочной средней (доли)

и генеральной дисперсией существует следующее отношение: дисперсия выборочной средней (доли) равна дисперсии признака в генеральной совокупности, деленной на число отобранных единиц, либо, что то же, на объем выборки:

$$\sigma_{\bar{x}}^2 = \frac{\sigma_0^2}{n}$$

и

$$\sigma_w^2 = \frac{p(1-p)}{n},$$

где σ_0^2 — генеральная дисперсия признака;
 p — доля признака в совокупности с альтернативными признаками;
 $p(1-p)$ — дисперсия альтернативного признака.

Корень квадратный из этих выражений носит название средней ошибки выборки. Так, средняя ошибка при выборочном определении средней

$$\sigma_{\bar{x}} = \sqrt{\frac{\sigma_0^2}{n}}. \quad (7.1)$$

Соответственно средняя ошибка при выборочном определении доли

$$\sigma_w = \sqrt{\frac{p(1-p)}{n}}. \quad (7.2)$$

Из формул (7.1) и (7.2) видно, что средняя ошибка выборки прямо пропорциональна среднему квадратическому отклонению признака. Поэтому чем больше колеблемость значений признака в генеральной совокупности, тем больше средняя ошибка выборки, и, наоборот, с уменьшением колеблемости значений признака уменьшается и размер возможной ошибки выборки.

Эти же формулы показывают, что средняя ошибка выборки обратно пропорциональна корню квадратному из числа наблюдений (объем выборки). Поэтому по мере возрастания объема выборки размер средней ошибки выборки уменьшается. Если, например, требуется уменьшить среднюю ошибку выборки в 2 раза, то для этого при прочих равных условиях необходимо увеличить объем выборки в 4 раза, чтобы уменьшить среднюю ошибку выборки в 3 раза, объем выборки следует увеличить в 9 раз и т. д.

Применение формул (7.1) и (7.2) предполагает, что известны генеральная дисперсия и генеральная доля. Однако в действительности эти показатели неизвестны, так как выборка для того и проводится, чтобы на основе выборочных показателей судить о неизвестных значениях соответствующих генеральных показателей. Ввиду этого возникает необходимость замены генеральной дисперсии (σ_0^2) и генеральной доли (p) другими, близкими к ним величинами. Такими величинами могут служить выборочная дисперсия (σ^2) и выборочная доля (ω).

В математической статистике доказывается, что математическое ожидание выборочной дисперсии собственно-случайной порционной выборки равно $\frac{n-1}{n} \sigma_0^2$ в случае определения средней признака и равно $\frac{n-1}{n} pq$, если находится доля. Здесь n — объем выборки, σ_0^2 — генеральная дисперсия, p — генеральная доля, а $q = 1 - p$. Откуда

$$\sigma_0^2 = \sigma^2 \frac{n}{n-1} \quad \text{и} \quad pq = \omega(1-\omega) \frac{n}{n-1}.$$

Нетрудно видеть, что при достаточно большом объеме выборки величина поправки $\frac{n}{n-1}$ существенного значения не имеет, так как знаменатель этой дроби незначительно отличается от числителя, а дробь в целом — от единицы. Например, если $n = 100$, то $\frac{n}{n-1} = \frac{100}{99} = 1,01$, если $n = 500$, то $\frac{n}{n-1} = \frac{500}{499} = 1,002$ и т. д. Следовательно, для выборки большого объема можно взять σ^2 вместо σ_0^2 и $\omega(1-\omega)$ вместо pq .

С учетом сказанного формулы средней ошибки выборки могут быть записаны так:

$$\sigma_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}} \quad (7.3)$$

$$\sigma_w = \sqrt{\frac{\omega(1-\omega)}{n}}. \quad (7.4)$$

Средняя ошибка выборки характеризует меру отклонений выборочной средней (или доли) от генеральной средней (или доли). При этом, как доказывается в математической статистике, с определенной вероятностью можно утверждать, что эти отклонения не превысят некоторую величину, которую можно назвать предельной ошибкой выборки. Обозначив предельную ошибку греческой буквой Δ (дельта), а буквой t — коэффициент доверия, зависящий от вероятности, с которой можно гарантировать, что предельная ошибка выборки не превысит t -кратную среднюю ошибку, можно написать следующее равенство:

$$\Delta = t \sigma. \quad (7.5)$$

Положение о том, что по своему абсолютному значению разность между выборочными и генеральными обобщающими показателями с определенной вероятностью не превысит предельной ошибки выборки, вытекает из существа закона больших чисел. Выдающемуся русскому математику П. Л. Чебышеву принадлежит следующая обобщающая формулировка закона больших чисел (принадлежит к выборке): с вероятностью, сколь угодно близкой

к единице, можно утверждать, что при достаточно большом объеме выборки и ограниченной генеральной дисперсии выборочные обобщающие показатели (средняя, доля) будут столь угодно мало отличаться от соответствующих генеральных обобщающих показателей.

Применительно к нахождению среднего значения признака теорема П. Л. Чебышева (с уточнениями А. М. Ляпунова) может быть записана так:

$$P[|\bar{x} - \bar{x}| \leq \Delta] = \Phi(t), \quad (7.6)$$

а для доли признака — соответственно

$$P[|w - p| \leq \Delta] = \Phi(t), \quad (7.7)$$

где

$$\Phi(t) = \frac{1}{2\pi} \int_{-t}^{+t} e^{-\frac{z^2}{2}} d(z).$$

Значения функции $\Phi(t)$ при различных значениях t определяются на основе специально составленных таблиц. Приведем некоторые из этих значений, применяемые наиболее часто:

t	1,0	1,96	2,0	2,58	3,0
Φ_t	0,683	0,950	0,954	0,990	0,997

Следовательно, предельная ошибка выборки отвечает на вопрос о точности выборки с определенной вероятностью, величина которой определяется значением коэффициента доверия t . Так, при $t = 1$ вероятность $\Phi(t)$ отклонения выборочных характеристик от генеральных на величину однократной средней ошибки равна 0,683. Это значит, что в среднем из каждой 1000 выборок 683 дадут обобщающие показатели (средняя, доля), которые будут отличаться от генеральных обобщающих показателей не более чем на величину однократной средней ошибки. При $t = 2$ вероятность $\Phi(t)$ равна 0,954, что означает, что из каждой 1000 выборок 954 дадут обобщающие показатели, которые будут отличаться от генеральных обобщающих показателей не более чем на двукратную среднюю ошибку выборки и т. д.

На основании теоремы Чебышева решается ряд задач выборочного наблюдения, в частности определение предельной ошибки выборки при заданной вероятности и определение численности выборки, необходимой для обеспечения заданной ее точности с определенной вероятностью. Рассмотрим это на примере 10%-ной собственно-случайной повторной выборки.

По приведенным выше данным о стаже 30 рабочих выборочная дисперсия составляет 54,87 и средняя ошибка выборки по формуле (7.3) равна:

$$\mu_x = \sqrt{\frac{54,87}{30}} = 1,35 \text{ (года)}.$$

Из формулы (7.6) следует, что предельные значения генеральной средней можно определить по формуле

$$\bar{x} = \bar{x} \pm \Delta.$$

Следовательно с вероятностью 0,683 ($t = 1$) можно утверждать, что средняя стаж всех рабочих цеха находится в пределах $10 \pm 1,35$, т. е. от 8,65 до 11,35 лет. С вероятностью 0,954 ($t = 2$) можно утверждать, что средний стаж всех рабочих цеха находится в пределах от 7,3 до 12,7 лет, и т. д.

Средняя ошибка выборочной доли по формуле (7.4) составляет

$$\mu_w = \sqrt{\frac{0,367 \times 0,633}{30}} = 0,088$$

0,087 — выборочная доля рабочих со стажем свыше 10 лет).

Следовательно, с вероятностью 0,683 ($t = 1$) можно утверждать, что доля рабочих цеха со стажем свыше 10 лет находится в пределах $0,367 \pm 0,088$, т. е. от 0,279 до 0,455, или от 27,9 до 45,5%. С вероятностью 0,954 ($t = 2$) можно утверждать, что доля рабочих цеха со стажем свыше 10 лет находится в пределах от 0,207 до 0,529 и т. д.

Определение необходимой численности выборки производится на основе алгебраического преобразования формулы (7.5).

Из этой формулы следует, что при определении средней

$$\Delta = t \mu_x = t \sqrt{\frac{\sigma^2}{n}}.$$

Отсюда

$$n = \frac{t^2 \sigma^2}{\Delta^2}. \quad (7.8)$$

Если выборка еще не производилась, а лишь планируется, то выборочная дисперсия, конечно, неизвестна. В таких случаях при использовании формулы (7.8) берутся ориентировочные значения σ^2 , полученные при ранее проведенных аналогичных обследованиях или на основе пробной выборки.

Обращаясь к рассматриваемому примеру, определим, какой должна быть объем выборки, чтобы с вероятностью 0,954 предельная ошибка выборки не превышала двух лет. По формуле (7.8) получаем:

$$n = \frac{2^2 \times 54,87}{0,2^2} = 55 \text{ человек}.$$

При определении доли имеем $\Delta = t \mu_w = t \sqrt{\frac{w(1-w)}{n}}$, откуда необходимая численность выборки равна:

$$n = \frac{t^2 w(1-w)}{\Delta^2}. \quad (7.9)$$

Вычислим по этой формуле численность выборки, которая необходима для того, чтобы с вероятностью 0,954 предельная ошибка выборки не превышала 0,15.

Подставляя в формулу условия задания, получаем величину необходимой численности выборки:

$$n = \frac{2^2 \times 0,367 \times 0,633}{0,15^2} = 41 \text{ человек.}$$

Из двух полученных значений численности выборки следует взять меньшую величину, т. е. 55 человек, ибо только в этом случае будет обеспечена заданная точность выборки как для средней, так и для доли, т. е. по обоим интересующим нас показателям.

Образим теперь выборочную совокупность способом собственно-случайного бесповторного отбора при том же объеме выборки (30 единиц, или 10% общей численности единиц). В этом случае так же как и при повторном отборе, можно применить жеребьевку, но каждый билет после записи проставленного на нем стажа в урну не возвращается.

Допустим, что в выборку попали рабочие, имеющие следующие табельные номера: 101, 90, 69, 194, 267, 270, 61, 287, 86, 70, 114, 276, 9, 32, 1, 30, 72, 74, 85, 237, 66, 12, 291, 258, 281, 148, 95, 16, 112, 27.

Эти рабочие имеют следующий стаж работы (лет): 3, 10, 4, 21, 23, 4, 1, 8, 5, 1, 1, 4, 12, 8, 30, 1, 11, 19, 2, 6, 10, 11, 6, 3, 5, 12, 4, 21, 5, 3. Сумма этих показателей составляет 256 лет, а средний стаж рабочих в выборочной совокупности равен: $x = 256 : 30 = 8,53$ (года). Так как в выборку попали девять рабочих со стажем свыше 10 лет, то выборочная доля составляет: $w = 9 : 30 = 0,30$, или 30%.

В математической статистике доказывается, что для бесповторного отбора формулы средней ошибки собственно-случайной выборки имеют следующий вид:

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}, \quad (7.10)$$

$$\mu_w = \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}, \quad (7.11)$$

где N — объем генеральной совокупности.

Сравнение формул (7.10) и (7.11) с формулами (7.3) и (7.4) показывает, что они отличаются лишь множителем $\left(1 - \frac{n}{N}\right)$, содержащимся под знаком корня. Так как этот множитель всегда меньше единицы, то ошибка выборки при бесповторном отборе

меньше, чем при повторном отборе. Для отбора необходимой численности единиц совокупности могут быть использованы и таблицы случайных чисел. См.: Велецкий И. Г. Князев Г. С. Основы математической статистики. М., Госстатиздат, 1960. Романовский В. И. Применение математической статистики в опшнм деле, 1947 и др.

величине ошибки выборки повторного отбора. По мере увеличения объема выборки множитель $1 - \frac{n}{N}$ стремится к нулю, а при $n = N$, т. е. когда будет подвергнута обследованию вся совокупность единиц, множитель $1 - \frac{n}{N}$ превращается в нуль и вместе с ней случайная ошибка репрезентативности становится равной нулю.

Следует отметить, что при бесповторном отборе часто применяются формулы ошибок повторного отбора. Эта замена находит свое оправдание для случаев, когда выборка составляет малую долю генеральной совокупности. Тогда отношение $n : N$ становится очень малой величиной и, следовательно, разность $1 - \frac{n}{N}$ будет мало отличаться от единицы.

Для рассматриваемого примера бесповторного отбора выборочная дисперсия составляет 53,77, откуда средняя ошибка выборки при определении стажа равна:

$$\mu_{\bar{x}} = \sqrt{\frac{53,77}{30} \left(1 - \frac{30}{300}\right)} = 1,27 \text{ (лет).}$$

Соответственно средняя ошибка доли равна:

$$\mu_w = \sqrt{\frac{0,3 \times 0,7}{30} \left(1 - \frac{30}{300}\right)} = 0,079.$$

Определение предельной ошибки выборки, т. е. отклонения фактного размера признака и доли в выборочной совокупности от той или другой стороны от генеральной средней и доли с заданной вероятностью, а также определение необходимой численности выборки производится аналогично тому, как это было показано применительно к повторному отбору.

Подставив, например, в формулу (7.5) значение средней ошибки $\mu_{\bar{x}}$, получим:

$$\Delta = t \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}.$$

Откуда необходимая численность выборки при определении средней равна:

$$n = \frac{t^2 \sigma^2 N}{\Delta^2 N + t^2 \sigma^2}. \quad (7.12)$$

При определении доли численности бесповторной выборки составляет:

$$n = \frac{t^2 w(1-w) N}{\Delta^2 N + t^2 w(1-w)}. \quad (7.13)$$

Механический отбор называется такая выборка, при которой отбор единиц производится механически, через определенный интервал.

При организации механического отбора единицы генеральной

совокупности предварительно располагают (обычно в списке) в каком-либо определенном порядке (по алфавиту, по географическому принципу, в порядке возрастания или убывания значения какого-либо показателя и т. д.), после чего отбирают заданное число единиц механически, через определенный интервал.

В тех случаях, когда единицы расположены в порядке, не связанном со значением изучаемого признака, величина интервала определяется как частное от деления объема генеральной совокупности на объем выборочной совокупности. Например, необходимо произвести 10%-ную механическую выборку рабочих сборочного цеха (см. табл. 7.1) для определения среднего стажа и доли рабочих со стажем более 10 лет. Объем выборки равен $\frac{300 \times 10}{100} = 30$ человек. Величина интервала составляет 10 человек (300 : 30).

Следовательно, в нашем примере, где имеется 30 бригад по 10 человек в каждой, необходимо отобрать по одному рабочему из каждой бригады. Допустим, что путем жеребьевки определено, что это будет сельмой рабочий. Тогда в выборочную совокупность войдут 30 рабочих, табельные номера которых будут: 7, 17, 27, 37, 47, 57, 67, 77, 87, 97, 107, 117, 127, 137, 147, 157, 167, 177, 187, 197, 207, 217, 227, 237, 247, 257, 267, 277, 287 и 297.

Эти рабочие имеют следующий общий производственный стаж (лет): 11, 10, 3, 1, 1, 28, 2, 7, 14, 19, 2, 4, 3, 19, 23, 2, 4, 9, 3, 4, 4, 3, 5, 6, 12, 11, 23, 11, 8 и 13. Средний стаж рабочих в выборочной совокупности составляет 8,83 года, дисперсия стажа — 48,9, а доля рабочих со стажем свыше 10 лет — 36,7%.

Широкое распространение в практике статистических работ имеет механическая выборка на основе предварительного расположения единиц генеральной совокупности по возрастающему (или убывающему) значению изучаемого или связанного с ним признака. В этих случаях величина интервала обычно определяется путем деления общего итога по какому-либо связанному с изучаемым показателем на число единиц, подлежащих отбору (см. «Обследования бюджетов семей трудящихся»).

Механический отбор имеет определенные преимущества перед собственно-случайной выборкой, так как дает обычно более близкое распределение отобранных единиц к распределению единиц в генеральной совокупности по изучаемым признакам, и выборка становится более репрезентативной. Кроме того, при механическом отборе проще организовать и легче проводить правильность отбора единиц.

Оценка точности выборки при механическом отборе производится по формулам собственно-случайной выборки. Это объясняется тем, что средняя ошибка выборки при механическом отборе меньше либо в крайнем случае равна средней ошибке собственно-случайной выборки. При этом более вероятно, что ошибка репрезентативности механической выборки будет не больше, чем ошибка собственно-случайной бесповторной выборки. Но для больших

чисел единиц в достоверности полученных результатов часто применяется формула ошибки собственно-случайной повторной выборки.

Для рассматриваемого примера средняя ошибка выборки составит:

а) для среднего стажа работы

$$\mu_x = \sqrt{\frac{48,9}{30}} = \sqrt{1,63} = 1,28 \text{ (года)},$$

б) для доли рабочих со стажем свыше 10 лет

$$\mu_p = \sqrt{\frac{0,367 \times 0,633}{30}} = 0,088.$$

Следовательно, с вероятностью 0,683 ($t = 1$) можно утверждать, что средний стаж всех рабочих цеха находится в пределах $8,83 \pm 1,28$, т. е. от 7,52 до 10,08 года. В отношении доли рабочих со стажем свыше 10 лет с вероятностью 0,683 ($t = 1$) можно утверждать, что она находится в пределах $0,367 \pm 0,088$, т. е. от 0,279 до 0,455, или от 27,9 до 45,5%.

В случае, когда генеральная совокупность неоднородна и это влияет на размер изучаемого признака, применяется предварительное деление ее на типически однородные группы (районы). Группировка производится по существенным признакам, которые связаны с изучаемыми признаками. Например, при изучении урожайности пахотных земель предварительно группируются по качеству почвы, при бюджетных обследованиях колхозы предварительно группируются по направлениям хозяйства и т. д. Затем отдельно из каждой такой типически однородной группы отбирают определенное число единиц либо механическим, либо собственно-случайным методом отбора. При этом общее число единиц выборочной совокупности обычно распределяется между типическими группами пропорционально удельному весу каждой группы в общей совокупности. Такой отбор носит название *пропорционального отбора*.

Возможен также и непропорциональный, так называемый *оптимальный отбор*, основанный на учете как удельного веса каждой группы в общей совокупности, так и варьирования признака по группам.

При типическом отборе в выборку попадают представители всех типических групп, поэтому достигается большая репрезентативность и большая точность выборки, чем при собственно-случайном или механическом отборе. При этом точность выборки зависит от того, насколько хорошо отобранные единицы репрезентуют среднюю (долю) каждой типической группы, а также в какой мере признаки, положенные в основу группировки по типам, действительно связаны с изучаемым.

Если например, из 300 рабочих сборочного цеха, по которому ранее определялся средний стаж работы, 150 человек имеют низкую квалификацию (1-й и 2-й разряды), 90 человек — среднюю

квалификацию (3-й и 4-й разряды) и 60 человек — высокую квалификацию (5-й и 6-й разряды) и если установлено, что квалификация рабочих тесно связана со стажем, то, выбрав при 10% выборке 30 человек путем случайного отбора или механического, рискуем получить недостаточно точную характеристику среднего стажа. Это может произойти в том случае, когда, допустив в выборку попадут главным образом рабочие низкой и средней квалификации и очень мало рабочих высокой квалификации или наоборот, когда в выборку попадет много рабочих высокой и средней квалификации и мало рабочих низкой квалификации.

Чтобы добиться в данном случае более точных результатов, поделим предварительно всех рабочих цеха (генеральная совокупность) на группы по квалификации, а затем уже из каждой такой типической однородной группы будем производить выборку необходимого числа рабочих способом механического отбора. При такой организации отбора выборочная совокупность будет лучше репрезентовать генеральную совокупность, так как отбор из отдельных типических групп обеспечивает пропорциональное представление в выборку рабочих разной квалификации.

По каждой типической группе составим список рабочих в порядке увеличения табельных номеров (порядковые номера рабочих в этих списках приведены в гр. 3 табл. 7.1). Внутри каждой типической группы произведем механическую группировку по десяткам и в каждом десятке отберем второго по счету рабочего.

Тогда из первой типической группы численностью в 150 человек в выборку попадут 15 рабочих, имеющих порядковые номера 2, 12, 22, 32, 42, 52, 62, 72, 82, 92, 102, 112, 122, 132 и 142.

Найдем для этой группы средний стаж и дисперсию стажа: $\Sigma x = 5 + 4 + 1 + 2 + 10 + 1 + 1 + 7 + 3 + 3 + 1 + 1 + 3 + 7 + 8 = 57$; $\Sigma x^2 = 25 + 16 + 1 + 4 + 100 + 1 + 1 + 49 + 9 + 9 + 1 + 1 + 49 + 49 + 64 = 339$.

Средний стаж $\bar{x}_1 = 57 : 15 = 3,8$ года; дисперсия $\sigma_1^2 = \frac{339}{15} - (3,8)^2 = 22,60 - 14,44 = 8,16$.

Так как в выборке нет ни одного рабочего со стажем свыше 10 лет, то выборочная доля по этому признаку равна нулю, точно так же нулю равна и дисперсия по этому признаку.

Из второй типической группы численностью в 90 человек в выборку попадут 9 рабочих, имеющих порядковые номера: 152, 162, 172, 182, 192, 202, 212, 222 и 232.

Для этой группы средний стаж составляет $\bar{x}_2 = 9,78$ года и дисперсия — $\sigma_2^2 = 17,95$.

Так как в выборку попало 4 рабочих со стажем свыше 10 лет, их доля в выборке равна $w_2 = 4 : 9 = 0,444$, или 44,4%, а дисперсия — $w_2(1 - w_2) = 0,444 \times 0,556 = 0,247$.

Наконец, из третьей типической группы численностью 60 человек в выборку попадут 6 рабочих, имеющих порядковые номера 242, 252, 262, 272, 282 и 292.

Для этой группы средний стаж составляет $\bar{x}_3 = 21,17$ года и дисперсия — $\sigma_3^2 = 25,66$.

Так как в выборке все 6 рабочих имеют стаж свыше 10 лет, то выборочная доля по этому признаку $w_3 = 6 : 6 = 1$, а дисперсия — $w_3(1 - w_3) = 1 \times 0 = 0$.

В целом средний стаж рабочих, вошедших в выборочную совокупность, составляет $\bar{x} = 9,07$ (года), а дисперсия $\sigma^2 = 57,9$.

Доля рабочих со стажем свыше 10 лет будет:

$$w = \frac{0 + 4 + 6}{15 + 9 + 6} = \frac{10}{30} = 0,33, \text{ или } 33\%,$$

а дисперсия по этому признаку $w(1 - w) = 0,33 \times 0,67 = 0,222$.

Расчет средней ошибки выборки при типическом отборе, пропорциональном численности единиц в типических группах, производится с помощью следующих формул:

	Повторная выборка	Бесповторная выборка
Для определения среднего размера изучаемого признака	$\mu_x = \sqrt{\frac{\sigma^2}{n}} \quad (7.14)$	$\mu_{\bar{x}} = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)} \quad (7.15)$
Для определения доли признака	$\mu_w = \sqrt{\frac{w(1-w)}{n}} \quad (7.16)$	$\mu_w = \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)} \quad (7.17)$

Сравнив формулы средней ошибки выборки типического отбора с соответствующими формулами собственно-случайного отбора, увидим, что они отличаются тем, что вместо дисперсий по всей выборочной совокупности [σ^2 и $w(1 - w)$] взяты средние величины из соответствующих показателей вариации по группам σ_1^2 и $w_1(1 - w_1)$.

Согласно закону сложения вариаций общая дисперсия σ^2 равна сумме межгрупповой дисперсии (δ^2) и средней величины из внутригрупповых дисперсий (σ^2): $\sigma^2 = \delta^2 + \sigma^2$.

Поэтому средняя ошибка типической выборки менее или в крайнем случае (если межгрупповая дисперсия равна нулю) равна средней ошибке собственно-случайной выборки. Отсюда вытекает, что выборку из генеральной совокупности, состоящей из единиц различающихся между собой групп, целесообразно производить путем типического отбора и нецелесообразно производить путем собственно-случайного отбора, при котором средняя ошибка выборки зависит от общей вариации признака.

Рассчитаем предельную ошибку выборки и найдем значение стандартной средней по данным типической выборки по математическим табл. 7.1. Выборочная средняя, как подсчитано выше,

составляет 9,07 года. Среднюю ошибку типической выборки при бесповторном отборе определим по формуле (7.15). Используя данные, приведенные на с. 212, найдем среднюю взвешенную дисперсию отдельных групп рабочих:

$$\bar{\sigma}^2 = \frac{8,16 \times 15 + 17,95 \times 9 + 25,8 \times 6}{15 + 9 + 6} = \frac{438,75}{30} = 14,62.$$

Тогда средняя ошибка типической выборки будет:

$$\mu_{\bar{x}} = \sqrt{\frac{14,62}{30} \left(1 - \frac{30}{300}\right)} = \sqrt{\frac{13,158}{30}} = \sqrt{0,4386} = 0,66 \text{ (года)}.$$

Произведем теперь сопоставление. Средняя ошибка при выборочном изучении величины стажа рабочих составила при собственно-случайном бесповторном отборе 1,27, а при типическом отборе — 0,66. Следовательно, применительно к данному примеру типическая выборка при прочих равных условиях дала более точные результаты по сравнению с собственно-случайным отбором. Уменьшение средней ошибки выборки при типическом отборе есть результат предварительного деления единиц генеральной совокупности на однородные группы по существенным признакам, которые связаны с изучаемыми признаками. Отсюда — важное значение правильного выбора группировочных признаков.

Уменьшение средней ошибки выборки влечет за собой уменьшение и предельной ошибки выборки. Так, с вероятностью 0,68 ($t=1$) можно утверждать, что средний стаж всех рабочих цеха находится в пределах $9,07 \pm 0,66$, или от 8,41 до 9,73 года. С вероятностью 0,954 ($t=2$) можно утверждать, что средний стаж всех рабочих цеха находится в пределах $9,07 \pm 1,32$, или от 7,75 до 10,39 года и т. д.

Аналогично обстоит и с предельной ошибкой доли. Выборочная доля рабочих со стажем свыше 10 лет, как подсчитано выше, составляет 0,33, или 33%. Среднюю ошибку доли определим по формуле (7.17). Используя данные, приведенные на с. 212, найдем среднюю взвешенную из дисперсий отдельных групп рабочих:

$$\bar{\omega}(1-\bar{\omega}) = \frac{0 \times 15 + 0,247 \times 9 + 0 \times 6}{15 + 9 + 6} = \frac{2,223}{30} = 0,074.$$

Тогда средняя ошибка доли при типическом отборе будет:

$$\mu_{\bar{\omega}} = \sqrt{\frac{0,074}{30} \left(1 - \frac{30}{300}\right)} = \sqrt{\frac{0,0666}{30}} = \sqrt{0,00222} = 0,047, \text{ или } 4,7\%.$$

Произведем и здесь сопоставление. Средняя ошибка при выборочном изучении доли рабочих со стажем свыше 10 лет составила при собственно-случайном бесповторном отборе 7,9%, а при типическом отборе — 4,7%. Следовательно, и при изучении доли признака типическая выборка при прочих равных условиях дает более точные результаты.

Типовой серийный отбор

В рассмотренных видах выборки (собственно-случайный, механический и типический отбор) речь шла об отборе отдельных единиц из генеральной совокупности. Но иногда целесообразно производить отбор не отдельных единиц, а целых групп их (гнезд, рядов) с тем, чтобы в таких группах подвергать наблюдению все без исключения единицы.

Например, при изучении производительности труда рабочих в той или иной отрасли народного хозяйства возможно организовать выборочное наблюдение так, чтобы отбирать не отдельных рабочих, а целые предприятия с тем, чтобы на этих предприятиях подвергать наблюдению всех без исключения рабочих. При различного рода выборочных обследованиях в сельском хозяйстве можно производить отбор целых колхозов с тем, чтобы в колхозах подвергать наблюдению все хозяйства колхозников и т. д. Серийный отбор применяется также при выборочном контроле за качеством продукции, при контрольных обходах, проводимых после зачета скота, и т. д.

Серийная выборка применяется в двух вариантах: 1) когда все серии имеют одинаковое количество единиц и 2) когда серии неодинаковы по объему. При этом наиболее распространенным и статистически более разработанным является первый вариант.

Преимущество серийного отбора перед отбором отдельных единиц в том, что его легче организовать. Но при серийном отборе значительно нарушается равномерность распределения отобранных единиц в пределах генеральной совокупности и более высокая ошибка выборки. Чтобы обеспечить ту же точность, что и при собственно-случайном индивидуальном отборе, требуется увеличить численность выборки.

Ошибки серийной выборки (с равными сериями) определяются по формулам, сходным с формулами ошибок собственно-случайной выборки. Допустим, что генеральная совокупность поделена на некоторое количество (S) равных по численности серий. Отвечает определенное число (s) серий в порядке случайной или механической выборки. При этом точность выборки будет зависеть от того, насколько хорошо средние показатели серий (\bar{x}_i) будут репрезентовать генеральную среднюю. Чем меньше средние показатели серий будут отклоняться от генеральной средней, тем точнее будут результаты выборки.

При серийном отборе (с равными сериями) мерой колеблемости является межсерийная выборочная дисперсия

$$\hat{\sigma}_x^2 = \frac{\sum (\bar{x}_i - \bar{\bar{x}})^2}{s} \quad (7.18)$$

при изучении среднего значения признака либо

$$\hat{\sigma}_{\bar{\omega}}^2 = \frac{\sum (\bar{\omega}_i - \bar{\bar{\omega}})^2}{s} \quad (7.19)$$

при изучении доли признака.

Здесь \bar{x}_i — серийные средние, \bar{x} — общая выборочная средняя, ω_i — доли изучаемого признака в сериях, $\bar{\omega}$ — общая выборочная доля и s — число серий, отобранных для обследования.

Средняя ошибка серийной выборки (μ) при равновеликих сериях определяется по таким формулам:

	Повторная выборка	Бесповторная выборка
При определении среднего размера изучаемого признака	$\mu_{\bar{x}} = \sqrt{\frac{\bar{x}_s^2}{s}} \quad (7.20)$	$\mu_{\bar{x}} = \sqrt{\frac{\bar{x}_s^2}{s} \cdot \frac{S-s}{S-1}} \quad (7.21)$
При определении доли данного признака	$\mu_w = \sqrt{\frac{\bar{\omega}_s^2}{s}} \quad (7.22)$	$\mu_w = \sqrt{\frac{\bar{\omega}_s^2}{s} \cdot \frac{S-s}{S-1}} \quad (7.23)$

Покажем расчет предельной ошибки выборки и определение границ генеральных средней и доли по данным серийной выборки на ранее приводившемся примере (см. табл. 7.1). Рабочие сборочного цеха поделены на 30 бригад по 10 человек в каждой. Для определения среднего производственного стажа рабочих цеха и доли рабочих со стажем более 10 лет путем 10%-ной выборки нужно на основе случайной жеребьевки выбрать три бригады. Пусть это оказались 10-я (табельные № 91—100), 24-я (табельные № 231—240) и 28-я (табельные № 271—280).

Определение границ генеральной средней достигается путем следующих расчетов. Средний стаж в 10-й бригаде равен $x_1 = 4,9$ года, в 24-й — $x_2 = 11,1$ года, в 28-й — $x_3 = 10,8$ года и в целом по выборке $\bar{x} = (4,9 + 11,1 + 10,8) : 3 = 8,9$ (года).

Межсерийная дисперсия составляет:

$$\bar{x}_s^2 = \frac{(4,9 - 8,9)^2 + (11,1 - 8,9)^2 + (10,8 - 8,9)^2}{3} = \frac{24,45}{3} = 8,15.$$

Для определения величины среднего стажа работы всех рабочих цеха нужно найти среднюю ошибку выборки. Так как отбор производился по схеме бесповторной выборки, то эта ошибка определяется по формуле (7.21):

$$\mu_{\bar{x}} = \sqrt{\frac{8,15}{3} \cdot \frac{(30-3)}{(30-1)}} = 1,59 \text{ (года)}.$$

Отсюда с вероятностью 0,683 ($t = 1$) можно утверждать, что средний стаж всех рабочих цеха находится в пределах $8,9 \pm 1,59$, или колеблется в пределах от 7,31 до 10,49 лет. С вероятностью 0,954 ($t = 2$) можно утверждать, что средний стаж рабочих цеха находится в пределах $8,9 \pm 3,18$, или от 5,72 до 12,08 года, и т. д.

Определение границ генеральной доли достигается путем следующих расчетов. Доли рабочих со стажем более 10 лет состав-

ляют в 10-й бригаде — 0,1, в 24-й — 0,4, в 28-й — 0,4 и в целом по выборке — 0,3. Межсерийная дисперсия равна:

$$\bar{\omega}_s^2 = \frac{(0,1 - 0,3)^2 + (0,4 - 0,3)^2 + (0,4 - 0,3)^2}{3} = \frac{0,06}{3} = 0,02.$$

Для определения доли рабочих со стажем более 10 лет по цеху в целом находим среднюю ошибку выборки по формуле (7.23):

$$\mu_w = \sqrt{\frac{0,02}{3} \cdot \frac{(30-3)}{(30-1)}} = \sqrt{0,0062} = 0,079, \text{ или } 7,9\%.$$

Следовательно, с вероятностью 0,683 ($t = 1$) можно утверждать, что доля рабочих со стажем более 10 лет по цеху в целом находится в пределах $30 \pm 7,9\%$, т. е. от 22,1% до 37,9%. С вероятностью 0,954 ($t = 2$) можно утверждать, что доля рабочих со стажем более 10 лет по цеху в целом находится в пределах $30 \pm 15,8\%$, т. е. от 14,2% до 45,8%, и т. д.

МАЛАЯ ВЫБОРКА

Понятие о малой выборке. Распространение выборочных характеристик (средняя, доля) на генеральную совокупность, основанное на действии закона больших чисел, предполагает достаточно большой объем выборки. Из теоремы П. Л. Чебышева вытекает, что чем больше объем выборки, тем выше репрезентативность выборочных характеристик. Наоборот, уменьшение объема выборки ухудшает репрезентативность, приводит к значительным ошибкам при определении генеральных характеристик.

В то же время не всегда возможен и целесообразен большой объем выборки. Изучение качества продукции, производительности труда, урожайности и других показателей в ряде случаев не позволяет производить наблюдение в больших масштабах. Возникает задача отыскания путей получения генеральных характеристик при малой выборке, объемом до 20 единиц. В математической статистике доказывается, что и при малых выборках возможно распространение выборочных характеристик на генеральную совокупность.

Новый разворот теории малой выборки положил английский математик-статистик Вильям Госсет, опубликовавший свои работы под псевдонимом Стьюдента. Затем эти работы продолжил английский математик-статистик Р. А. Фишер. Большой вклад в теорию малой выборки сделан советскими математиками В. Н. Романовским и А. Н. Колмогоровым.

Граници и предельная ошибка малой выборки

В предыдущем параграфе было показано, что выборочная дисперсия определяется по формуле

$$s^2 = \frac{\sum (x - \bar{x})^2}{n}$$

и что она связана с генеральной дисперсией равенством

$$\sigma_0^2 = \sigma^2 \frac{n}{n-1}.$$

При этом поправочный коэффициент $\frac{n}{n-1}$ для малой выборки в отличие от обычной имеет существенное значение, и при расчете средней ошибки малой выборки вместо неизвестной генеральной дисперсии (σ_0^2) используется выборочная дисперсия, умноженная на этот поправочный коэффициент. Поэтому средняя ошибка малой выборки ($\mu_{м.в}$) определяется по следующей формуле:

$$\mu_{м.в} = \sqrt{\frac{\sigma_{м.в}^2}{n}}, \quad (7.24)$$

где $\sigma_{м.в}^2$ — дисперсия в малой выборке, исчисляемая по формуле

$$\sigma_{м.в}^2 = \frac{\sum (x - \bar{x})^2}{n-1}, \quad (7.25)$$

а предельная ошибка малой выборки ($\Delta_{м.в}$) — по формуле

$$\Delta_{м.в} = t \mu_{м.в}. \quad (7.26)$$

Вероятности, соответствующие отдельным значениям коэффициента доверия t , неодинаковы при различных объемах малой выборки (n). Значения этих вероятностей при различных t и n приводятся в специальных таблицах (таблицы вероятностей по распределению Стьюдента). Так, например, вероятность того, что предельная ошибка малой выборки не превзойдет t -кратную среднюю ее ошибку $S(t) = P(|x - \bar{x}| \leq \Delta_{м.в})$, равна:

$t \backslash n$	5	10	15
1	0,626	0,657	0,666
2	0,884	0,923	0,935
3	0,960	0,985	0,990

Исходя из предположения о нормальном распределении признака в генеральной совокупности Стьюдент установил, что для малой выборки величина $t = \frac{x - \bar{x}}{\mu_{м.в}}$ имеет следующий закон распределения:

$$S(t) = C \left(1 - \frac{t^2}{n-1}\right)^{-\frac{n}{2}},$$

где постоянная C зависит только от числа n — объема выборки.

По мере увеличения объема выборки (n) это распределение приближается к нормальному и при $n > 20$ оно уже мало отличается от нормального.

При этом вероятности того, что величина t не превзойдет некоторое число t , т. е. что $-t_1 < t < +t_1$, определяется при помощи интеграла Стьюдента:

$$P_n(|t| < t_1) = \int_{-t_1}^{+t_1} S(t) dt,$$

значения которого приводятся в специальных таблицах. Применяются и другие варианты этого интеграла.

С помощью интеграла Стьюдента решается ряд задач малой выборки, в частности определение интервала, в котором находится генеральная средняя.

Допустим, что из 300 рабочих цеха (см. табл. 7.1) в случайном порядке отобрано 15 человек, имеющих следующие табельные номера (взяты табельные номера первых пятнадцати рабочих со стр. 200): 116, 54, 40, 177, 12, 37, 145, 42, 127, 84, 205, 271, 226, 256 и 297.

Произведем исчисление среднего стажа этих рабочих и дисперсии:

5, 7, 4, 9, 11, 1, 18, 18, 3, 10, 6, 22, 13, 10, 13. Итого 150.
25, 49, 16, 81, 121, 1, 324, 324, 9, 100, 36, 484, 169, 100, 169. Итого 2008.

Рассчитав по этим данным средний стаж и дисперсию, получим:

$$\bar{x} = 10 \text{ лет}, \quad \sigma_{м.в}^2 = 36,3.$$

Находим далее среднюю квадратическую ошибку малой выборки:

$$\mu_{м.в} = \sqrt{\frac{36,3}{15}} = 1,56.$$

Определим с вероятностью 0,935 границы интервала, в котором находится генеральная средняя. По таблице находим значение t при $n = 15$ и $S(t) = 0,935$. Оно равно 2. Следовательно, генеральная средняя, в данном случае средний стаж всех рабочих цеха, равна: $\bar{x} \pm t \mu_{м.в} = 10 \pm 2 \times 1,56 = 10 \pm 3,12$, т. е. средний стаж всех рабочих с вероятностью 0,935 находится в пределах от 6,9 до 13,1 года.

При проведении малой выборки нужно иметь в виду, что определение предельной ошибки ее возможно лишь при условии, что распределение признака в генеральной совокупности является нормальным или близким к нему.

Бесспорно, что в ряде случаев при выборочном определении среднего значения признака целесообразна организация малой

выборки. Однако по точности результатов и их надежности выборка большего объема имеет преимущество перед малой выборкой.

4. МОМЕНТНО-ВЫБОРОЧНОЕ НАБЛЮДЕНИЕ

Понятие о моментно-выборочном наблюдении

Особой разновидностью выборочного наблюдения является так называемое моментно-выборочное, или, сокращенно, моментное наблюдение. Сущность его заключается в фиксации наличия отдельных элементов изучаемого процесса на определенные моменты времени, которые являются составными частями всего периода обследования, последовательные и непрерывно следующие друг за другом, без учета продолжительности данного элемента.

Например, при изучении структуры рабочего времени рабочего какого-либо производственного участка выделяются два элемента «работа» и «простой». В течение рабочей смены через определенные интервалы времени (о порядке их определения см. ниже) производится обход рабочих и фиксация по состоянию на данный момент времени выделенных элементов рабочего времени («работа» или «простой») без учета продолжительности каждого из этих элементов. Допустим, что за семичасовую смену каждого было произведено 84 наблюдения за работой рабочего. При этом в 74 случаях рабочий работал (элемент «работа») и в 10 случаях — не работал (элемент «простой»). Исходя из этих данных определяем структуру рабочего времени. Удельный вес элемента «работа» составляет $88\% \left(\frac{74}{84} \times 100 \right)$ и элемента «простой» — $12\% \left(\frac{10}{84} \times 100 \right)$. В общей сумме сменного времени элемент «работа» составляет 370 мин $\left(\frac{420 \times 88}{100} \right)$ и элемент «простой» — 50 мин $\left(\frac{420 \times 12}{100} \right)$. Точно так же по данным моментной выборки определяются и средняя структура рабочего времени всех рабочих производственного участка.

Эта разновидность выборочного наблюдения с успехом применяется при изучении структуры рабочего времени и использования производственного оборудования в промышленности, покупок и продаж в торговле, использования вагонного парка на железнодорожном транспорте и во многих других случаях, когда элементы изучаемого процесса следуют во времени один за другим последовательно и непрерывно и исключена возможность их одномоментного осуществления. Резко снижая трудоемкость изучения подобных массовых процессов, метод моментных наблюдений дает возможное увеличение объема выборки исследуемых единиц (при тех же расходах на проведение обследования), а это в свою очередь значительно повышает степень репрезентативности собираемой информации.

Генеральная и выборочная совокупности при моментной выборке

Цель моментной выборки, как и обычной, — репрезентативная оценка показателей генеральной совокупности на основе обследования выборочной совокупности. Однако при моментной выборке понятия генеральной и выборочной совокупностей, а также единицы совокупности имеют свое, особое содержание. Генеральная совокупность в этом случае представляет полную продолжительность времени, в течение которого производится наблюдение изучаемого процесса. Например, при изучении структуры рабочего времени генеральная совокупность представляет рабочую смену продолжительностью 8 часов (либо 7 часов и т. д.). Выборочная же совокупность представляет собой совокупность отдельных моментных состояний объекта, отобранных из генеральной совокупности. Эти моментные состояния объекта, представляющие отдельные элементы изучаемого процесса на определенные моменты, выступают как единицы совокупности и как единицы наблюдения.

Способы отбора, обеспечивающие репрезентативность моментной выборки

Для образования выборочной совокупности при моментной выборке используется собственно-случайный, механический и плановый отбор. Возможно и сочетание этих способов в целях достижения более высокой репрезентативности. В отличие от обычной выборки все перечисленные способы отбора организуются только в виде бесповторной выборки, так как однажды отобранные моменты не могут быть отобраны повторно.

Собственно-случайной называется такая выборка, при которой отбор моментов из всего изучаемого времени производится независимо, в случайном порядке. Для этого используется либо жеребьевка, либо таблицы случайных чисел (см. гл. VIII).

Механическая — это такая выборка, при которой все изучаемое время (генеральная совокупность) предварительно делится механически на нужное число равных по продолжительности отрезков времени, в конце которых производится регистрация единиц (элементов). При этом, как и при обычной выборке, под механическим делением изучаемого времени на интервалы понимается такое деление, которое не связано с какими-нибудь систематическими особенностями исследуемого объекта, от которого может зависеть исследуемый признак. Это необходимо для предотвращения систематической ошибки репрезентативности механического отбора.

Плановый отбор называется такая выборка, которая обеспечивает равномерное (планомерное) попадание в выборку моментов от всех характерных частей генеральной совокупности. Тем самым плановый отбор обеспечивает наивысшую репрезентативность моментной выборки.

Как и при обычной выборке, оценка репрезентативности выборки, в данном случае выборочной доли, определяется с помощью показателей средней и предельной ошибок выборочной доли, исчисляемых по формулам бесповторного отбора.

5. РАСПРОСТРАНЕНИЕ ВЫБОРОЧНЫХ ДАННЫХ НА ГЕНЕРАЛЬНУЮ СОВОКУПНОСТЬ

Результаты всякого выборочного обследования должны быть распространены на всю генеральную совокупность. Это конечная практическая цель всякого выборочного наблюдения.

Расчет объемных показателей генеральной совокупности на основании данных выборочного наблюдения называется в статистике *распространением выборочных характеристик на всю совокупность*. Существуют два способа такого распространения: 1) *способ прямого пересчета* и 2) *способ коэффициентов*.

В первом случае средний размер признака, найденный в результате выборочного обследования, умножается на число единиц генеральной совокупности. Установив, например, в результате выборочного обследования продуктивности скота, находящегося в личной собственности, что средний годовой настриг шерсти с одной овцы составил 2,5 кг, и зная, что всего в районе в личной собственности имеется 1000 овец, можно получить величину валового настрига шерсти: $2,5 \times 1000 = 2500$ кг, или 2,5 т. Если при этом известно, что средняя ошибка выборки с той или иной вероятностью равна $\pm 0,1$ кг и что, следовательно, генеральная средняя с той же вероятностью колеблется в пределах от 2,4 до 2,6 кг, то общий валовой настриг шерсти будет колебаться от 2,4 до 2,6 т.

Во втором случае в данные сплошного обследования вносятся соответствующие поправки, относительный размер которых устанавливается по результатам выборочного обследования. Эти поправки вносятся на основании сопоставления данных сплошного и выборочного обследований.

Способ коэффициентов применяется при контрольных обходах по проверке численности скота, находящегося в личной собственности

населения, проживающего в сельской местности. Если в результате контрольного обхода выявляется недоучет скота, сравнительно с данными хозяйственных книг, то производится соответствующее уточнение этих данных.

Известно, что в результате контрольных обходов по району выявлены данные, расходящиеся с данными хозяйственных книг. Тогда производится расчет поправочных коэффициентов следующим образом (табл. 7.2).

Генеральную районную сводку итогов учета скота могут быть внесены следующие поправки, уточняющие численность поголовья.

Таблица 7.3

РАСЧЕТ ЧИСЛЕННОСТИ ПОГОЛОВЬЯ

	Численность поголовья	
	зарегистрированного в хозяйственных книгах	с поправкой на недоучет
Коровы	5 600	$5600 \times 1,005 = 5628$
Прочий крупный рогатый скот	5 000	$5000 \times 1,01 = 5050$

ПРАКТИКА ПРИМЕНЕНИЯ ВЫБОРОЧНОГО НАБЛЮДЕНИЯ

Выборочное наблюдение в революционной статистике

Применение выборочного наблюдения в русской статистике насчитывает не одну сотню лет. Известно, что начало применения примитивных форм выборочного наблюдения в России относится еще к XVII—XVIII вв. Например, боярин Б. Н. Морозов (XVII в.) составил своим приказам грамоту (инструкцию) по выделению специальных учетных пунктов, уполномоченных производить выборочный пробный обмолот по трем категориям хлеба (хорошего, среднего и плохого) и о результатах сообщать хозяину фольварку в виде отчета (записки в «опытные» книги). Затем этот метод получил все большее распространение, особенно при изучении крестьянского хозяйства.

Значительного совершенствования выборочное наблюдение достигает в работе русской земской статистики. Ставя перед собой задачу всестороннего изучения крестьянского хозяйства пореформенной России и будущим образом в видах и средствах для проведения статистических исследований, земская статистика пошла по пути широкого применения выборочного метода. Изучение статистики велось по пути проведения крестьянских хозяйств Вятской губернии 1884 г., повторное обследование Вятской губернии в 1901—1902 гг., быстрое обследование крестьянских хозяйств Тамбовской губернии в 1906 г., оценочно-статистическое исследование Пензенской губернии в 1911 г. и др.

Земские статистики использовали в своих работах различные способы отбора, в том числе и типичный. Заслугой земских статистиков является также сочетание сплошных подворных переписей по краткой программе с выборочными обследованиями по более широкой и детальной программе.

РАСЧЕТ ПОПРАВОЧНЫХ КОЭФФИЦИЕНТОВ

	Число голов в контрольных точках		Поправочный коэффициент (стр. 7.2, стр. 1)
	данные хозяйственных книг	установлено при контрольных обходах	
А	1	2	3
Коровы	600	603	1,005
Прочий крупный рогатый скот	500	505	1,01

Выборочные наблюдения советской статистики

Советская статистика с первых шагов своего существования обращается к выборочному методу для изучения состояния промышленности и сельского хозяйства, также для обследования бюджетов отдельных групп населения. В настоящее время выборочный метод имеет широкое распространение и применяется при исследовании большого круга вопросов в различных отраслях народного хозяйства и культуры.

В первую очередь следует сказать о применении выборочного метода в переписях населения. Так, при проведении переписи населения в 1970 г. по ряду показателей (о месте работы населения и продолжительности этой работы в 1969 г., занятии, общественной группе переписываемых) информация была получена выборочным путем. Была применена механическая 25%-ная выборка жилых помещений.

С помощью выборочного метода изучаются бюджеты рабочих, служащих и колхозников. СССР имеет самую большую и наиболее полную сеть выборочных обследований по изучению уровня жизни населения, которые дают ценные сведения о материальном благосостоянии и культурном уровне трудящихся нашей страны.

На основе выборочных обследований изучается большой круг вопросов работы промышленности: межотраслевые связи, качество продукции, использование вновь введенных производственных площадей и мощностей, использование оборудования и рабочей силы, производительность труда, состояние заработной платы и т. д.

Выборочный метод успешно применяется при изучении многих вопросов экономики сельского хозяйства: полноты учета урожая хлопка в колхозах, размера потерь зерновых при уборке, всхожести различных семян, продуктивности скота, находящегося в личной собственности колхозников, размеров и структуры посевных площадей и урожайности на приусадебных участках колхозников и др.

Выборочный метод применяется и при изучении других отраслей народного хозяйства. Так, путем выборочных обследований собираются данные о ценах продажи сельскохозяйственных продуктов на колхозном рынке и объеме продажи этих продуктов.

Применяется выборочный метод в самых разнообразных формах: выборка объектов для наблюдения (например, в бюджетных обследованиях), выборочная разработка данных сплошного наблюдения (например, данных переписей населения), выборка во времени, которая применяется и для проверки достоверности данных, полученных в результате сплошного наблюдения (определение износа основных фондов, учет численности скота). Дальнейшее развитие советской статистики связано с резким увеличением объема выборочных обследований. Успешная разработка теории выборочного метода и все более широкое использование математического и экономико-статистических исследований являются важной пред-

посылкой расширения сферы применения выборочного исследования. А наличие вычислительной техники (в том числе и электронной) позволяет советской статистике осуществлять выборочные обследования в больших масштабах. Рассмотрим методику организации одной из наиболее распространенных статистических работ — проводимых выборочным методом.

Обследования бюджетов семей колхозников

Одной из наиболее важных статистических работ, осуществляемых выборочным методом, является обследование бюджетов семей рабочих, служащих и колхозников. В СССР органами государственной статистики обследуется более 50 тыс. семей рабочих, служащих, инженерно-технических работников и колхозников. Основной задачей государственной статистики СССР в области бюджетных обследований является получение данных, необходимых для изучения и планирования роста материального благосостояния населения.

При бюджетных обследованиях семей колхозников выборочная代表性 создается путем двухступенчатого отбора. Сначала производится отбор колхозов, а затем в отобранных колхозах выделяются семьи колхозников для обследования их бюджетов.

Отбор колхозов производится статистическими управлениями области (края, республики) в результате комбинации типичной и механической выборки. В качестве типичных признаков, по которым колхозы объединяются в группы, берутся производственное направление колхозов и распределенный доход в расчете на один человеко-день. Все колхозы области (края, республики) разбиваются на группы по производственному направлению: зерновые, животноводческие, производящие технические культуры и т. д.

В пределах каждой группы колхозов по производственному направлению составляется список всех колхозов в порядке убывания размера распределенного на один человеко-день дохода. Это обеспечивает дальнейшую типизацию при последующем механическом отборе. В списке по каждому колхозу указываются его наименование, размер распределенного дохода на один человеко-день, число дворов в нем, нарастающий итог числа дворов и название района, в котором находится колхоз. Отбор колхозов из этого списка производится механически через определенный интервал по нарастающему итогу числа дворов. Величину интервала определяют путем деления всего числа дворов колхозников данной области (края, республики) на число колхозов, которое необходимо отобрать.

В разных республиках для обследования отбирается различное число семей в пределах каждого отобранного колхоза. В частности, в Украинской ССР отбирается по 25 семей в каждом отобранном колхозе. Для иллюстрации техники отбора предположим, что в какой-либо области нужно обследовать 550 семей колхозников, причем в пределах одного колхоза должно быть обследо-

доваю 25 семей. Допустим, далее, что в этой области имеется 220 022 колхозных двора. Чтобы определить величину интервала, найдем сначала число колхозов, в которых должны вестись бюджетные обследования. Это число определяется путем деления числа семей, подлежащих обследованию, на число хозяйств колхозников, обследуемых в пределах одного колхоза. Для условия нашего примера оно составит 22 колхоза ($550 : 25$). Отсюда величина интервала, через который нужно отбирать колхозы, будет равна 10 001 ($220\,022 : 22$).

В каждой группе по производственному направлению первоначально отбирается колхоз, в котором находится двор, имеющий в списке номер 5001 (середина интервала), затем колхозы, в которых находятся хозяйства, имеющие номера 15 002, 25 003 и т. д. Инструкция по обследованию бюджетов колхозников предусматривает, что в каждом районе, как правило, должен обследоваться один колхоз. Поэтому если в процессе отбора в выборку попадет второй колхоз того же района, то такой колхоз заменяется колхозом из другого района, соседним по списку.

После того как колхозы отобраны, производят проверку на репрезентативности, т. е. проверку того, насколько точно отобранные колхозы характеризуют все колхозы области (края, республики). Репрезентативность отобранной совокупности колхозов устанавливается путем сопоставления ряда показателей по всем колхозам области (края, республики) и по отобранным колхозам.

Сопоставляются, в частности, такие показатели: 1) в расчете на один колхозный двор: число человеко-дней, на которые распределен доход; поголовье скота (по видам); уборочная площадь (по группам культур); сумма денег, предназначенных к выдаче (по группам культур); объем денег, предназначенных к выдаче (в стоимостном выражении), и др.; 2) в расчете на один человеко-день: размер выданных зерновых и бобовых, денег и ряд других показателей. Если по перечисленным показателям отобранные колхозы значительно отличаются от генеральной совокупности (всех колхозов области), то колхозы с сильно отличающимися показателями заменяют другими колхозами, показатели которых более близки к показателям генеральной совокупности.

После того как получена репрезентативная выборочная совокупность колхозов, производят отбор семей колхозников путем комбинирования типической и механической выборки. В каждом отобранном колхозе все хозяйства делятся на две группы: имеющие и не имеющие коров. В пределах каждой из этих групп хозяйства записываются в список в порядке убывания количества отработанных человеко-дней. Это обеспечивает дальнейшую типизацию отбора — попадание в выборку хозяйства с различным количеством отработанных человеко-дней. Отбор производится отдельно из каждой группы пропорционально числу хозяйств в ней в порядке механической выборки, т. е. через определенный интервал. Величина интервала определяется путем деления числа всех хозяйств колхозников на число хозяйств, которое необходимо

иметь для обследования бюджетов. Если, например, в колхозе 125 хозяйств и из них 400 имеют коров, а 125 не имеют и должно быть отобрано 25 хозяйств, то величина интервала будет $125 : 25 = 5$ хозяйств. В первой группе надо отобрать $400 : 25 = 16$ хозяйств, а во второй — $125 : 25 = 5$ хозяйств. В выборку в другую группу попадут хозяйства, имеющие в списке номера 11 (середина интервала из 21 хозяйства), 32, 53 и т. д.

Если в число отобранных попадет семья, которая в своем составе не имеет трудоспособных, то она заменяется другой семьей по списку. Обследование отобранных хозяйств колхозников производится при условии согласия семьи на ведение бюджетных записей. Если одна семья отказывается от обследования, то она заменяется другой, состоящей из такого же числа лиц и соответствующей первой по следующим признакам: характер выполняемой в колхозе работы, количество начисленных за год человеко-дней, обеспеченность скотом в личной собственности.

После отбора хозяйств колхозников проверяется их репрезентативность путем сопоставления средних показателей отобранных хозяйств с аналогичными средними показателями по всем хозяйствам колхозников данной области (края). К этим показателям относятся:

- а) число всех лиц в семье, в том числе наличных членов семьи, работающих в государственных и кооперативных организациях;
- б) количество начисленных всей семье человеко-дней;
- в) общая площадь приусадебного посева, в том числе:
 - а) картофеля, б) овощей и бахчевых;
- г) число голов скота в личной собственности колхозников:
 - а) крупного рогатого скота; б) овец и коз; в) свиней.

В случае значительной нерепрезентативности хозяйств колхозников производится частичная замена их из числа хозяйств отобранных колхозов.

СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ
ВЗАИМОСВЯЗЕЙ1. ВЗАИМОСВЯЗИ ОБЩЕСТВЕННЫХ ЯВЛЕНИЙ
И ЗАДАЧА ИХ СТАТИСТИЧЕСКОГО ИЗУЧЕНИЯ

Изучение взаимосвязи социально-экономических явлений — одна из важнейших задач статистики

Советская статистика при изучении общественных явлений исходит из положений исторического материализма о всеобщей связи явлений и процессов общественной жизни. Эта связь выражается в том, что все стороны жизни общества, как единицы социального организма, взаимодействуют между собой. Так, сложные и многообразные связи существуют между экономическими и политическими процессами, между различными отраслями народного хозяйства, между производством и потреблением продукции и т. п. Например, сельское хозяйство нуждается в машинах, удобрениях и т. д., которые поставляет ему промышленность; в свою очередь промышленность получает от сельского хозяйства сырье.

Взаимосвязи между общественными явлениями, их признаки а следовательно, и между статистическими показателями обнаруживаются как в целом для народного хозяйства, так и для отдельных его сторон, отраслей, предприятий и т. п. Так, рост производительности труда на заводе при прочих равных условиях ведет к снижению себестоимости единицы продукции и увеличению прибыли. Это позволяет выделять больше средств в фонды развития производства, материального стимулирования и социально-культурных мероприятий, что в свою очередь способствует дальнейшему росту производительности труда. Наоборот, повышение, например, текучести кадров работников предприятия ведет к снижению уровня производительности труда, повышению себестоимости продукции, уменьшению прибыли.

Связи между общественными явлениями, отдельными их признаками очень многообразны и сложны. Однако в каждом конкретном случае одни признаки влияют на другие и обуславливают их изменение, т. е. выступают как признаки-факторы (фак-

торные признаки), другие же признаки зависят от факторных признаков, изменяются под влиянием их изменения, т. е. выступают как результаты действия этих факторов, как результативные признаки.

Статистика, изучающая социально-экономические явления, не может ограничиться изучением отдельно взятого явления, изолированно от тех связей, которые оно имеет с другими явлениями. Наоборот, она стремится по возможности охватить весь комплекс взаимосвязанных явлений, чтобы, статистически изучив их, дать им связь с числовое выражение.

Связь между признаками может иметь различный характер. В одних случаях значения результативного признака целиком и полностью определяются значениями признаков-факторов, так что каждой определенной системе значений (вариантов) факторных признаков соответствует одно (или несколько) строго определенное значение результативного признака. Такие связи (зависимости) называются функциональными. Функциональные связи имеют место в математике, физике, астрономии. Так, длина окружности полностью определяется величиной ее радиуса, а площадь треугольника — длиной его сторон. Имеют место такие связи и в общественных явлениях, в частности в экономике. Так, если при определенной оплате труда дневная ставка установлена в размере S руб., то начисленная работнику за отработанные дни (T) сумма заработной платы (S) будет функционально связана с числом отработанных дней: $S = ST$. Или другой пример: общая выручка за проданный товар определенного вида и сорта целиком определяется его количеством и ценой единицы товара. Если куплен товар двух сортов, на каждый из которых установлена определенная цена (x_1 и x_2), то средняя цена единицы купленного товара будет функционально зависеть от удельных весов I и II сорта в общем объеме покупки ($\bar{x} = x_1\omega_1 + x_2\omega_2$, где ω — удельные веса).

При функциональной связи нам обычно известен полный перечень факторов, от которых зависит результативный признак, рассматриваемый в определенной связи, известен, как правило, и механизм взаимосвязи в виде того или иного уравнения, функции. Функциональная связь сохраняет при этом свою силу и проявляется в каждом отдельном случае наблюдения, для каждой отдельной единицы данной совокупности. Функциональные связи изучаются в статистике с помощью индексного метода (см. гл. X).

Однако в области массовых социально-экономических явлений эта или иная количественная закономерность чаще проявляется иначе — как зависимость распределения значений результативного признака от значений признаков-факторов. Такого рода связи (зависимости) называются стохастическими. Так, если распределение работников по полу неодинаково в разных отраслях

промышленности, то можно говорить о стохастической связи между этими двумя признаками: состав работников по полу статистически зависит от отрасли промышленности. Другой пример. Пусть совокупность колхозов разбита на несколько групп в зависимости от количества удобрений, внесенных в расчете на 1 га посева определенной культуры. Разобьем каждую такую группу на подгруппы по уровню урожайности этой культуры. Если распределение колхозов по урожайности будет различным в разных группах, то по количеству внесенных удобрений, т. е. будет изменяться при переходе от группы к группе, то между количеством внесенных удобрений (в расчете на 1 га) и урожайностью имеет место стохастическая связь.

Частным случаем стохастической связи является корреляционная связь. При этой связи одному и тому же значению признака-фактора могут соответствовать в отдельных случаях (у отдельных единиц совокупности) самые различные значения результативного признака, так что с изменением признака-фактора *меняется среднее значение результативного признака*. Следовательно, при корреляционной связи между значениями признаков нет строгого и точного соответствия в каждом отдельном случае, у каждой единицы совокупности, а наблюдается лишь известное соотношение, корреляция.

Так, например, между урожайностью и количеством внесенных удобрений есть, конечно, прямая связь. Однако если в нескольких колхозах на каждый гектар внесено одинаковое количество удобрений, то нельзя, разумеется, ожидать, что во всех этих колхозах и урожайность тоже будет одинаковой. Или если в колхозе А удобрений внесено больше, чем в колхозе Б, то это еще не означает, что урожайность в А будет обязательно больше, чем в Б. Это объясняется тем, что кроме количества удобрений урожайность зависит также от очень большого числа других факторов (качество семян и почвы, сроки сева и уборки, сроки и качество проведения различных агротехнических мероприятий и др.), а сочетание различных значений других факторов в каждом колхозе может быть самым различным. Вполне может оказаться поэтому, что в колхозе, где больше внесено удобрений, урожайность будет ниже, чем в другом колхозе, где при меньшем количестве удобрений лучше почва и качество ее обработки, выше качество семян и т. п.

Аналогичный характер носит, например, зависимость между уровнем производительности труда и стажем работы. При прочих равных условиях увеличение производственного стажа (до определенного предела) должно сопровождаться повышением выработки за единицу времени, так как по мере увеличения стажа совершенствуются навыки в работе, опыт работника. Однако в каждом отдельном рабочего уровень производительности труда зависит также от многих других факторов и обстоятельств (качество сырья и материалов, состояние оборудования и т. п.), влияние которых взаимно переплетается.

Наличие множества причин, обстоятельств и факторов, которые действуют одновременно и во взаимной связи, причем мера, степень влияния каждого из них на результативный признак непостоянно неизвестна и изменяется в зависимости от конкретных сочетания, — таковы отличительные особенности корреляционных связей. Ввиду этого корреляционная связь между результативным признаком и тем или иным признаком-фактором является только в *общем и среднем*, при прочих равных условиях, когда влияние других факторов, не являющихся объектом исследования, устранено путем их осреднения. В соответствии с этим больших чисел такое осреднение достигается при достаточном большом числе единиц совокупности, числе наблюдений.

При корреляционной связи, в отличие от функциональной, *точно неизвестен ни полный перечень всех признаков-факторов, влияющих на результативный признак, ни точный механизм их взаимодействия с ним* в виде той или иной математической формулы, функции.

Корреляционная связь является «неполной», «нестрогой».

Связи прямые и обратные. Как функциональные, так и корреляционные связи по их направлению могут быть прямыми и обратными. Если направление изменения результативного признака совпадает с направлением изменения признака-фактора, то связь называется прямой. Следовательно, при прямой связи с увеличением значений признака-фактора результативный признак тоже увеличивается, а при уменьшении — уменьшается. Примерами таких связей могут служить уже рассмотренные связи между стажем работы и производительностью труда, между количеством удобрений и урожайностью. Нужно, однако, иметь в виду, что эти связи являются прямыми лишь при изменении факторных признаков в определенных пределах. Так, производительность труда работника повышается с увеличением стажа работы, но когда стаж работы достигает определенной величины, на уровне производительности труда начинает сказываться возраст и ее повышение постепенно замедляется, а затем прекращается и может смениться снижением. Аналогично обстоит дело и с повышением урожайности при увеличении количества удобрений, внесенных на 1 га: рост урожайности в этом случае также не будет беспредельным, и когда количество удобрений возрастет до определенного, оптимального в данных условиях уровня, урожайность достигнет максимума и при дальнейшем увеличении количества удобрений расти не будет и может даже снижаться из-за перенасыщения почвы.

Если направление изменения результативного признака не совпадает с направлением изменения признака-фактора, то связь между ними называется обратной. Так, при прочих равных условиях с ростом производительности труда снижается себестоимость единицы продукции, а снижение себестоимости ведет к повышению рентабельности производства.

Связи
прямолинейные
и криволинейные,
одnofакторные и
многофакторные

По аналитическому выражению (формуле) связи могут быть прямолинейными и криволинейными. Если зависимость результативного признака от данного признака фактора может быть выражена уравнением прямой линии, то связь называется прямой линией (линейной), если же зависимость выражается уравнением какой-либо кривой (гиперболы, параболы и т. п.), то связь называется криволинейной.

Функциональные связи точно выражаются тем или иным аналитическим уравнением, корреляционные же связи могут быть выражены при помощи аналитического уравнения лишь приближенно.

Если исследуется зависимость результативного признака только от одного признака-фактора, то связь называется одnofакторной. Если при этом связь является функциональной, то это означает, что результативный признак зависит только от данного фактора. Если же связь является корреляционной, то включение в аналитическое уравнение только одного фактора означает, что от влияния других, прочих факторов мы абстрагируемся, элиминируем, устраняем их действие. Такая корреляция называется парной, так как при этом рассматриваются только два признака (пара признаков).

Если же исследуется корреляционная зависимость результативного признака одновременно от нескольких признаков-факторов, то связь называется многофакторной, а корреляция — множественной (совокупной).

2. ОСНОВНЫЕ СТАТИСТИЧЕСКИЕ МЕТОДЫ И ПРИЕМЫ ИЗУЧЕНИЯ СТОХАСТИЧЕСКИХ ВЗАИМОСВЯЗЕЙ

При исследовании различных взаимосвязей между общественными явлениями статистика опирается на их теоретический и качественный анализ. Для подтверждения его выводов, для выявления и количественной характеристики взаимосвязей статистика использует свои специфические приемы и методы. Для исследования функциональных связей применяется балансовый метод (гл. IV) и индексный метод (гл. X).

Для изучения стохастических связей используются метод параллельных рядов, метод аналитических группировок, дисперсионный анализ и анализ корреляций и регрессий.

К числу наиболее простых и распространенных методов анализа взаимосвязанных явлений относится метод сопоставления параллельных рядов. Сущность его заключается в том, что полученные в результате сводки и обработки материалы располагают параллельными рядами либо по признаку пространства, либо по признаку времени. Совместное изучение та-

ких рядов дает возможность проследить соотношение и направление изменений сопоставляемых признаков.

Для использования метода сравнения параллельных рядов в статистическом изучении связей необходим анализ сопоставляемых рядов, установление наличия между ними причинных связей, а также тесного сопутствия. Например, только потому, что между урожайностью и себестоимостью продукции сельского хозяйства причинная связь, становится возможным построение, а также и сопоставление параллельных рядов этих показателей.

Таблица 8.1

УРОЖАЙНОСТЬ И СЕБЕСТОИМОСТЬ ЗЕРНА И САХАРНОЙ СВЕКЛЫ
В РАЙОНАХ ДВУХ СОВРЕМЕНЮЩИХСЯ РАЙОНОВ

	Зерно (без кукурузы)		Сахарная свекла	
	район А	район Б	район А	район Б
Урожайность — ц с 1 га	26	21	250	280
Себестоимость 1 ц — руб.	3,2	3,6	2,0	1,8

Из таблицы видно, что между себестоимостью и урожайностью данной культуры имеется обратная связь: там, где выше урожайность, себестоимость более низкая.

Если сравниваемые ряды состоят из большого числа единиц, то направление связи для разных единиц может оказаться различным. Пусть, например, имеются следующие данные по 10 колхозам района.

Таблица 8.2

РАСЧЕТ КОЭФФИЦИЕНТА ФЕХНЕРА И КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ РАНГОВ

Урожайность зерна (ц с 1 га)	Себестоимость 1 ц зерна (руб.) у	Знак отклонения		Совпаде- ние (+) или несовпаде- ние (-) знаков	Ранги		d	d ²
		от \bar{x}	от \bar{y}		по x	по y		
3,1	+	—	Н	1	2	8	—7	49
2,9	+	—	Н	2	9,5	—7,5	55,25	
2,9	+	—	Н	3	9,5	—6,5	42,25	
3,5	+	—	С	4	5	—1	1	
3,2	+	—	С	5	7	—2	4	
3,3	+	—	С	6	6	0	0	
3,7	—	+	Н	7	3	4	16	
3,8	—	+	Н	8	2	6	36	
3,6	—	+	Н	9	4	5	25	
4,0	—	+	Н	10	1	9	81	
								310,5
$\bar{x} = 3,4$								

Как правило, чем выше урожайность, тем ниже себестоимость, однако здесь есть и исключения (см., например, колхозы 1 и 2 в т. п.).

Для ориентировочного выявления наличия связи и ее направления в подобных случаях можно найти для каждого ряда средние (арифметическую невзвешенную) и определить, в какую сторону отклоняются от нее значения признака у каждой единицы. Если отклонения в одном ряду достаточно часто совпадают по направлению (знаку) с отклонениями в другом ряду, то можно говорить о прямой связи, если же значительно чаще встречаются противоположные по направлению (знаку) отклонения, то связь является обратной. Если совпадение и несовпадение знаков отклонения встречается приблизительно одинаково часто, то связь либо отсутствует, либо очень слабо выражена.

Для приближенного определения направления связи и грубой оценки ее тесноты может быть использован коэффициент Фехнера:

$$K = \frac{C - H}{C + H}, \quad (8.1)$$

где C — число случаев совпадения знаков отклонений,
 H — число случаев несовпадения этих знаков.

Этот коэффициент принимает значения в пределах от +1 (знаки всех отклонений совпадают, связь прямая) до -1 (знаки всех отклонений не совпадают, связь обратная). При $K = 0$ связи отсутствует или очень слабая.

В нашем примере $K = (2-8) : 10 = -0,6$, т. е. между урожайностью и себестоимостью существует довольно тесная обратная связь.

Другой показатель, который может быть использован для ориентировочного выяснения наличия и направления корреляционной связи, называется коэффициентом корреляции рангов. Если коэффициент Фехнера основан на согласованности знаков отклонений от средней, то коэффициент ранговой корреляции учитывает согласованность рангов, т. е. мест, занимаемых единицами, расположенными в порядке возрастания или убывания величины признаков. Если места (ранги), занимаемые отдельными единицами в ранжированном ряду по одному признаку, точно соответствуют их местам в ранжированном ряду по другому признаку, то можно говорить о тесной прямой связи между этими признаками. Наоборот, если возрастание одного признака сопровождается непрерывным убыванием другого признака, так что ранги единиц изменяются в обратном порядке, то между признаками имеется тесная обратная связь. Наконец, при отсутствии той или иной согласованности в изменении рангов связь между признаками либо отсутствует, либо очень слабая.

Если разность между рангами по одному и другому признаку

значить через d , а число единиц в ряду через n , то коэффициент корреляции рангов выразится так:

$$r = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}. \quad (8.2)$$

Если для каждой единицы ранг по одному признаку равен рангу по другому признаку, то все $d = 0$, а $r = 1$ (тесная прямая связь). Если же первому месту по величине одного признака соответствует последнее место по величине другого, второму месту — первое с конца и т. д., то $r = -1$ (тесная обратная связь). Чем ближе связь, тем r ближе к +1 или к -1. При r близком к нулю связь слабая или отсутствует вообще.

В нашем примере (табл. 8.2) колхозы расположены в порядке убывания урожайности, так что их порядковые номера означают ранги по этому признаку. Там же показаны и ранги по величине себестоимости, причем у второго и третьего по порядку колхозов этот средний ранг, так как в ранжированном по величине себестоимости ряду они занимают 9 и 10 места.

В данном случае $r = 1 - \frac{6 \times 310,5}{10 \times (100 - 1)} = -0,88$, что подтверждает прежний вывод о достаточно тесной обратной связи.

Коэффициент корреляции рангов удобен тем, что он может быть применен и в тех случаях, когда невозможно точно измерить значения признаков, хотя ранжировать единицы можно (например, расположив студентов в порядке успеваемости по математике и статистике и т. п.).

Метод аналитических группировок

К числу важнейших приемов исследования взаимосвязей относится метод аналитических группировок (метод группировок в сочетании с методом обобщающих показателей). Чтобы выявить зависимость с помощью этого метода, можно произвести группировку единиц совокупности по признаку-фактору и для каждой группы вычислить среднюю или относительную величину резульативного признака. Сопоставляя затем изменение резульативного признака по мере изменения признака-фактора, можно выявить направление и характер связи между ними.

Замечательные образцы применения статистических группировок и групповых средних для установления связей и взаимозависимостей общественных явлений дают такие ленинские работы, как «Развитие капитализма в России», «Новые данные о законах развития капитализма в земледелии» и многие другие. В статье «Наша цифра», анализируя размеры зарплаток рабочих до и после революции 1905 г., В. И. Ленин с помощью статистических группировок показывает, что средняя заработная плата рабочих на крупных фабриках увеличилась после революции более значительно, чем на мелких. Объясняется это тем, что «...более энергичная и дружная стачечная борьба рабочих на крупных фабриках

привела в результате к более высокому увеличению заработной платы¹.

В книге «Развитие капитализма в России», рассматривая развитие горной промышленности в Донбассе, В. И. Ленин проследил зависимость между производительностью труда горняков и размером шахт по количеству рабочих в них и приходит к выводу, что производительность труда повышается вместе с увеличением размеров шахт, даже независимо от применения машин².

В работе «Новые данные о законах развития капитализма в земледелии» на основе обширного материала американской статистики В. И. Ленин с помощью аналитической группировки показывает, что по мере увеличения стоимости продуктов, производимых фермой, наблюдается правильное повышение интенсивности земледелия³.

Таблица 8.4

КЛАССИФИКАЦИЯ РАБОЧИХ, УРОВЕНЬ МЕХАНИЗАЦИИ ИХ ТРУДА
И ВЛИЯНИЕ ДНЕВНОЙ ВЫРАБОТКИ НА ФОРМОВКУ БЕТОНА
ОДНОМУ ИЗ ПРЕДПРИЯТИЙ

Разряд рабочего	Уровень механизации труда (процентов)	Дневная выработка (м³)	№ п/п	Разряд рабочего	Уровень механизации труда (процентов)	Дневная выработка (м³)
1	2	3	4	1	2	3
2	35	3,0	26	3	69	5,0
3	59	6,5	27	2	48	2,5
3	44	4,8	28	4	82	6,8
3	55	5,7	29	4	98	6,6
2	39	2,8	30	3	63	6,3
3	56	4,7	31	4	79	7,9
2	78	4,2	32	3	41	4,6
4	44	5,3	33	3	45	4,2
3	43	2,0	34	2	75	4,8
3	76	6,5	35	3	45	5,8
3	58	5,1	36	4	51	4,9
4	41	5,5	37	3	55	4,3
2	49	3,0	38	4	95	6,4
2	58	3,6	39	4	90	7,1
4	58	4,5	40	4	70	7,1
4	61	6,7	41	3	56	4,4
3	42	5,6	42	3	57	5,1
3	46	5,2	43	2	48	5,0
2	35	3,2	44	3	72	6,1
4	55	5,4	45	3	52	5,9
3	38	4,5	46	2	33	3,8
3	35	5,5	47	3	55	4,6
3	25	2,5	48	2	30	3,4
4	50	6,2	49	2	67	5,5
2	47	4,1	50	3	57	5,9

Таблица 8.3

Приходится на 1 акр земли долларов

Группы ферм по стоимости продукта	Расхода на наемн. труд	Расхода на удобрение	Стоимости всего снота	Стоимости орудий и машин
0	0,08	0,01	2,97	0,19
1— 50 долл.	0,06	0,01	1,78	0,38
50— 100 »	0,08	0,03	2,01	0,48
100— 250 »	0,11	0,05	2,46	0,62
250— 500 »	0,19	0,07	3,00	0,82
500—1000 »	0,36	0,07	3,75	1,07
1000—2500 »	0,67	0,08	4,63	1,21
2500 и более »	0,72	0,06	3,98	0,72

Статистические группировки и групповые средние, как важнейший прием исследования взаимосвязи, находят широкое применение в работах советских экономистов. Они применяются при решении таких вопросов, как экономическая эффективность специализации производства, зависимость между производительностью труда и себестоимостью продукции, между размерами предприятий по величине основных производственных фондов и производительности труда и т. д.

Покажем использование аналитических группировок и групповых средних для анализа связи на следующем примере (см. табл. 8.4).

У отдельных рабочих наблюдаются различные соотношения между квалификацией (разрядом) и производительностью труда (дневной выработкой). В одних случаях повышению разряда соответствует более высокая выработка, в других же, наоборот, несмотря на более высокий разряд, выработка оказывается ниже (см., например, 1-го и 9-го рабочих, 2-го и 8-го и т. д.). Различные соотношения наблюдаются у отдельных рабочих также между уровнем механизации труда и выработкой, даже при одинаковой квалификации. В большинстве случаев у рабочих одинаковой квали-

фикации с повышением уровня механизации труда повышается и дневная выработка. Однако есть и исключения. Так 1-й и 5-й рабочие имеют одинаковый разряд, причем несколько лучше механизирован труд у 5-го рабочего. Выработка же, наоборот, выше у 1-го рабочего.

Все это свидетельствует о том, что если между рассматриваемыми признаками есть связь, то она является корреляционной.

Для выяснения наличия связи между уровнем механизации труда и его производительностью (дневной выработкой) произведем группировку рабочих по проценту механизации труда и выведем для каждой группы среднюю дневную выработку. Выведем также и средний разряд рабочего в каждой группе.

Составляя изменение средней выработки с изменением уровня механизации труда, мы видим, что по мере повышения уровня механизации увеличивается и средняя дневная выработка, что является, конечно, вполне закономерным. Из данных таблицы

¹ Ленин В. И. Полн. собр. соч., т. 23, с. 429.

² См. там же, т. 3, с. 491—493.

³ Там же, т. 27, с. 190.

Таблица 85
ЗАВИСИМОСТЬ ПРОИЗВОДИТЕЛЬНОСТИ ТРУДА ОТ УРОВНЯ ЕГО МЕХАНИЗАЦИИ

Уровень механизации труда (процентов)	Число рабочих	Средний разряд	Средняя дневная выработка на одного рабочего (м³)
Менее 45	14	2,8	4,0
45—64	23	3,0	4,9
65 и выше	13	3,3	6,2
Итого	50	3,0	5,0

видно также, что с увеличением уровня механизации труда повышается и средний разряд рабочих: у 14 рабочих, труд которых механизирован менее чем на 45%, средний разряд ниже, чем у остальных рабочих, тогда как у 13 рабочих с наиболее высоким уровнем механизации труда наиболее высок и средний разряд. Это также вполне закономерное явление, так как работу с применением машин и механизмов поручают в первую очередь более квалифицированным рабочим.

Квалификация рабочих, однако, сама по себе влияет на уровень производительности труда, так как более квалифицированные рабочие имеют лучшую подготовку, навыки к работе, больший производственный опыт. Поэтому представляет интерес проследить зависимость дневной выработки и от этого фактора.

Таблица 86
ЗАВИСИМОСТЬ ПРОИЗВОДИТЕЛЬНОСТИ ТРУДА ОТ КВАЛИФИКАЦИИ РАБОЧИХ

Разряд рабочих	Число рабочих	Средний процент механизации труда	Средняя дневная выработка (м³)
2	13	49	3,8
3	24	52	5,0
4	13	70	6,2
Итого	50	56	5,0

При сопоставлении данных табл. 85 и 86 возникает вопрос: можно ли считать, что повышение средней выработки объясняется в каждом случае только увеличением величины группировочного признака? Ведь наряду с этим повышается величина и второго признака, который также влияет на выработку: так, в табл. 85 более высокая выработка 13 рабочих с наиболее высоким уровнем механизации труда одновременно сопровождается и более высоким разрядом (3,3), а в табл. 86 наиболее высокой выработке 13 рабочих 4-го разряда соответствует и наиболее высокий уровень механизации труда (70%). Очевидно, что поскольку производительность труда зависит как от уровня механизации труда, так и

квалификации рабочих, постольку изменение средней выработки в обоих случаях объясняется изменением обоих этих факторов.

И в частности, наиболее высокая выработка 13 рабочих последней группы в табл. 85 (6,2 м³) объясняется не только тем, что их труд является наиболее механизированным (65% и выше), но и тем, что они имеют в среднем и более высокую квалификацию (средний разряд 3,3). Соответственно наиболее высокая выработка 13 рабочих 4-го разряда в табл. 86 (6,2 м³) объясняется не только их более высокой квалификацией, но и тем, что у них наиболее высок в среднем уровень механизации труда (70%).

Чтобы выявить влияние каждого из этих факторов в более или менее «чистом» виде, нужно устранить, элиминировать влияние другого фактора, используя для этого комбинационный группировку по обоим факторам.

Таблица 87

ЗАВИСИМОСТЬ ПРОИЗВОДИТЕЛЬНОСТИ ТРУДА ОТ УРОВНЯ ЕГО МЕХАНИЗАЦИИ И ОТ КВАЛИФИКАЦИИ РАБОЧИХ

Уровень механизации труда (процентов)	Число рабочих				Средняя дневная выработка (м³)			
	всего	в том числе по разрядам			группы рабочих	в том числе рабочих по разрядам		
		2-й	3-й	4-й		2-й	3-й	4-й
Менее 45	14	5	7	2	4,0	3,2	4,2	5,4
45—64	23	5	14	4	4,9	3,6	5,3	5,4
65 и выше	13	3	3	7	6,2	4,8	5,9	6,9
Итого	50	13	24	13	5,0	3,8	5,0	6,2

Каждая из трех последних граф таблицы показывает изменение выработки в зависимости от роста уровня механизации труда рабочих соответствующего разряда, т. е. влияние квалификации на выработку здесь элиминировано. Во всех трех группах рабочих по квалификации по мере повышения уровня механизации труда растет и выработка. Исключение представляют лишь подгруппы рабочих 4-го разряда с уровнем механизации труда менее 45% и 45—64%, где выработка оказалась одинаковой. Это может объясняться случайными обстоятельствами, так как число рабочих в этих подгруппах невелико (2 и 4 человека).

При последних показателях каждой строки таблицы характеризует зависимость выработки от квалификации при более или менее полном элиминировании уровня механизации труда. Однако элиминирование влияния этого фактора является неполным, так как интервалы группировки довольно широки. Поэтому показатели выработки в каждой строке показывают ее изменение в за-

висимости от квалификации не при полностью одинаковом уровне механизации труда, а лишь при более или менее одинаковом уровне. Если, например, средний уровень механизации труда у 14 рабочих 1 группы составляет 37,5%, то у 5 рабочих 2-го разряда он равен только 34,2%, а у двух рабочих 4-го разряда — 42,5%. Неполнота элиминирования объясняется как взаимосвязью уровня механизации и квалификации, так и небольшим числом рабочих в каждой подгруппе.

Значительная наглядность в анализе связи достигается с помощью графиков. Для графического изображения зависимости результативного показателя (y) от одного фактора (x) при элиминировании влияния другого фактора (z) могут быть использо-

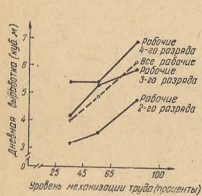


Рис. 8.1. Зависимость средней дневной выработки от уровня механизации труда

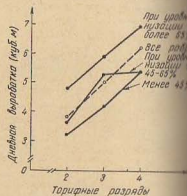


Рис. 8.2. Зависимость средней дневной выработки от квалификации (тарифных разрядов) рабочих

ваны линейные диаграммы в виде ломаных кривых. В прямоугольной системе координат по оси абсцисс откладывается тот фактор, влияние которого нужно изобразить (x), а по оси ординат откладываются значения результативного показателя. Для каждой группы по x и подгруппы по z , т. е. для каждой клетки комбинационной таблицы, исчисляются средние значения результативного показателя (y_{xz}), которые наносятся на график в виде точек, абсциссами которых являются середины интервалов фактора x . Соединив отрезками прямой последовательные значения y_{xz} по каждой подгруппе, получим для каждой из них ломаную кривую, выражающую зависимость y от фактора x при элиминировании фактора z .

Данные табл. 8.7 изображены графически на рис. 8.1 и 8.2. Аналитическая группировка позволяет не только подтверждать наличие связи и ее направление, но и измерить ее тесноту. Для числовой характеристики тесноты связи могут быть использованы

показатели вариации результативного признака: общая его дисперсия σ^2 и межгрупповая дисперсия

$$\sigma^2 = \frac{\sum (\bar{y}_i - \bar{y})^2 m_i}{\sum m_i},$$

где \bar{y}_i — групповые средние,

m_i — числа единиц в группах.

Известно, что группировка произведена так, что признак-фактор в пределах каждой группы совсем не варьирует (например, тарифная ставка рабочих по тарифным разрядам и т. п.), группировки по конкретному признаку без объединения отдельных его значений в одну группу, т. е. без образования интервалов). Если при этом число единиц в каждой группе достаточно велико, то в силу закона больших чисел влияние на y прочих факторов будет усреднено, так что вариация y будет определяться только вариацией x . Если результативный признак y совсем не связан с x , то групповые средние \bar{y}_i не будут изменяться с изменением x , т. е. будут равны одна другой и равны общей средней \bar{y} , а межгрупповая дисперсия σ^2 будет равна нулю.

Наоборот, если результативный признак y функционально связан с признаком-фактором x , то в каждой группе внутригрупповая дисперсия σ_i^2 будет равна нулю, так как признак x внутри группы не варьирует. Средняя из групповых дисперсий σ^2 тоже будет равна нулю, и согласно правилу сложения дисперсий межгрупповая дисперсия σ^2 совпадет с общей дисперсией σ^2 .

В промежуточных случаях, т. е. при корреляционной связи между y и x , межгрупповая дисперсия будет находиться между нулем и σ^2 , характеризую вариацию y , объясняемую вариацией x . Поэтому в качестве показателя тесноты связи между y и x при аналитической группировке может быть использовано отношение межгрупповой дисперсии к общей дисперсии результативного признака, называемое эмпирическим коэффициентом детерминации:

$$\eta_x^2 = \frac{\sigma^2}{\sigma^2}. \quad (8.3)$$

Коэффициент детерминации (η_x^2) показывает, какая часть общей вариации (дисперсии) результативного признака y объясняется вариацией группировочного признака-фактора x . Остальная же часть общей вариации y объясняется вариацией прочих факторов. При отсутствии связи коэффициент детерминации равен нулю, а при функциональной связи — единице.

Фактически, однако, в большинстве случаев дело осложняется двумя обстоятельствами: 1) тем, что при группировке по непрерывному или дискретному признаку образуются интервалы, так что внутри групп остается вариация признака, и связанная с ней вариация результативного признака; 2) тем, что признак-фактор x может быть корреляционно связан с одним или несколькими из прочих факторов, связанными в свою очередь с y . Вследствие первого обстоятельства межгрупповая дисперсия в таких случаях отражает не всю ва-

риацию, объяснимую фактором x , а только часть ее. Вследствие же того, что отстоятельства с изменением от группы к группе признака x будут изменяться и корреляционно связанные с ним другие факторы, так что межгрупповая дисперсия будет объясняться не только вариацией x , но и вариацией связанных с ним факторов.

Обычно в качестве показателя тесноты связи используют также эмпирическое корреляционное отношение — корень квадратный из коэффициента детерминации ($\eta = \frac{\delta}{\sigma}$). Как и η^2 , η находится от 0 до 1: чем ближе оно к 1, тем теснее связь.

Вычислим эмпирический коэффициент детерминации по данным табл. 8.4, используя в качестве группировочного признака фактор разряд рабочих (см. табл. 8.6). Общую дисперсию выработки y определим по формуле (6.20), т. е. как разность между средним квадратом и квадратом средней (расчет Σy^2 приведен ниже в табл. 8.12):

$$\sigma_y^2 = \bar{y}^2 - (\bar{y})^2 = \frac{1337,96}{50} - 5^2 = 1,7592.$$

Расчеты, необходимые для вычисления межгрупповой дисперсии, показаны в следующей таблице.

Таблица 8.4

К РАСЧЕТУ МЕЖГРУППОВОЙ ДИСПЕРСИИ

Разряд рабочих	Число рабочих m_i	Средняя дневная выработка (\bar{y}_i)	$\bar{y}_i - \bar{y}$	$(\bar{y}_i - \bar{y})^2$	$(\bar{y}_i - \bar{y})^2 m_i$
2-й	13	3,76	-1,24	1,5376	19,9888
3-й	24	5,03	0,03	0,0009	0,0216
4-й	13	6,18	1,18	1,3924	18,1012
Итого	50	5,00			38,1116

$$\delta^2 = \frac{\Sigma (\bar{y}_i - \bar{y})^2 m_i}{\Sigma m_i} = \frac{38,1116}{50} = 0,7622.$$

Следовательно, эмпирический коэффициент детерминации и корреляционное отношение равны:

$$\eta^2 = \frac{\delta^2}{\sigma^2} = \frac{0,7622}{1,7592} = 0,4333;$$

$$\eta = \sqrt{0,4333} = 0,658.$$

Таким образом, более 43% общей вариации выработки объясняется вариацией тарифных разрядов. Связь между этими признаками достаточно тесная.

Аналогично может быть измерена теснота связи для отдельных групп комбинационной таблицы, т. е. при более или менее полном элиминировании одного из группировочных признаков.

Отметим, что при группировке с образованием интервалов величина эмпирического коэффициента детерминации и корреляционного отношения зависит

от ширины групп и их границ. Так, по данным табл. 8.7 межгрупповая дисперсия $\delta^2 = 0,10$, $\eta^2 = 0,347$. Если же при группировке рабочих по уровню механизации труда образовать 4 группы (менее 40, 40—60, 60—80, 80—100%), то $\delta^2 = 0,419$, $\eta^2 = 0,462$. Расхождение между этими двумя значениями η^2 объясняется тем, что в первом случае вариация уровня механизации внутри групп уменьшается, а соответственно больше и объясняемая ею часть вариации y , которая уменьшается в δ^2 .

Рассматривая исходные данные как выборку из гипотетической генеральной совокупности, можно оценить значимость полученного коэффициента детерминации и корреляционного отношения с помощью дисперсионного F -критерия Фактора рассматриваемого в курсе математической статистики.

Комбинационная группировка может быть использована также для изучения стохастической связи между атрибутивными признаками. Так, В. И. Ленин, анализируя сборы, поступившие в годы подъема рабочего движения перед первой мировой империалистической войной на «правдивские» и «ликвидаторские» газеты, составил такую таблицу:

Таблица 8.9

Приходится сборов:	Из каждого рубля сборов в газеты:	
	правдивстов	ликвидаторов
от рабочих	87 копеек	44 копейки
не от рабочих	13 копеек	56 копеек
Всего	1 р. 00 копеек	1 р. 00 копеек

Связь между социальной принадлежностью лиц, делавших взносы, и направлением газет выражается здесь в том, что рас-

Таблица 8.10

Определение студентов по отраслевым группам вузов и по полу на начало 1971/72 учебного года)

Отраслевая группа вузов	Всего студентов (тыс. человек)	в том числе доля (часть)	
		мужчины	женщины
	m_i	w_{1i}	w_{2i}
Промышленность, строительство, транспорт и связь	2 088	0,62	0,38
Сельское хозяйство	423	0,69	0,31
Экономика и право	344	0,40	0,60
Здоровоохранение, физкультура и спорт	330	0,44	0,56
Проведение, искусство и киноматерия	1 412	0,33	0,67
Итого	4 597	0,51	0,49

Ленин В. И. Полн. собр. соч., т. 25, с. 232.

пределение сборов на «правдистские» газеты значительно отличается от распределения сборов на газеты «ликвидаторов». В то время как марксистские газеты опирались на рабочий класс, ликвидаторские газеты имели слабую связь с рабочим классом, отражали интересы буржуазии.

Может быть измерена и теснота связи между атрибутивными признаками. Рассмотрим, например, следующие данные (табл. 8.10 на стр. 243).

Если бы распределение студентов по полу совершенно не зависело бы от отрасли вуза, то доли мужчин и женщин во всех отраслевых группах были бы одинаковы и совпадали с их долями в общей численности студентов ($w_1 = 0,51$; $w_2 = 0,49$). Фактически этого нет, так что, очевидно, связь между отраслью вуза и составом студентов по полу имеется. Для измерения ее тесноты можно было использовать коэффициент взаимной сопряженности Л. А. Чупрова:

$$C = \sqrt{\frac{\chi^2}{(k_1 - 1)(k_2 - 1) \sum m_{ij}}} \quad (8.11)$$

где

$$\chi^2 = \sum_{i=1}^{k_1} m_i \sum_{j=1}^{k_2} \frac{(w_{ij} - w_{.j})^2}{w_{.j}},$$

где i — номер отраслевой группы вузов (I признак);

j — номер группы по полу (II признак);

k — число групп по соответствующему признаку;

w_{ij} — частость подгруппы i в группе j ;

$w_{.j}$ — частость группы j во всей совокупности;

m_i — число единиц в группе i .

При независимости признаков $w_{ij} = w_{.j}$, $\chi^2 = 0$, $C = 0$. При функциональной связи $C = 1$. В нашем примере:

$$\begin{aligned} \chi^2 &= 2088 \times \left[\frac{(0,62 - 0,51)^2}{0,51} + \frac{(0,38 - 0,49)^2}{0,49} \right] + \\ &+ 423 \times \left[\frac{(0,69 - 0,51)^2}{0,51} + \frac{(0,31 - 0,49)^2}{0,49} \right] + 344 \times \\ &\times \left[\frac{(0,40 - 0,51)^2}{0,51} + \frac{(0,60 - 0,49)^2}{0,49} \right] + \\ &+ 330 \times \left[\frac{(0,44 - 0,51)^2}{0,51} + \frac{(0,56 - 0,49)^2}{0,49} \right] + \\ &+ 1412 \times \left[\frac{(0,33 - 0,51)^2}{0,51} + \frac{(0,67 - 0,49)^2}{0,49} \right] = 362,14, \\ \text{откуда } C &= \sqrt{\frac{362,14}{4597 \times (5-1) \times (2-1)}} = 0,140. \end{aligned}$$

Следовательно, между отраслью вуза и полом студентов есть определенная связь, но ее теснота невелика.

Частным случаем связи атрибутивных признаков является связь количественных признаков, каждый из которых имеет два взаимозаключающих варианта. В этом случае комбинационная таблица имеет такой вид:

Таблица 8.11

Способность выполнения плана сдачи в эксплуатацию жилья при переходе бригад на новую форму организации труда

	Число бригад		
	выполнивших план	не выполнивших плана	всего
Бригады, перешедшие на новую форму организации труда	$a = 49$	$b = 1$	$a + b = 50$
Бригады, не перешедшие на новую форму организации труда	$c = 41$	$d = 9$	$c + d = 50$
Итого	$a + c = 90$	$b + d = 10$	100

Для измерения тесноты связи в этом случае используется коэффициент ассоциации (равный коэффициенту С):

$$A = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} = \frac{49 \times 9 - 41 \times 1}{\sqrt{50 \times 50 \times 90 \times 10}} = 0,267. \quad (8.5)$$

Следовательно, между переходом на новую форму организации труда и выполнением плана есть связь, переход на новую форму организации труда способствует выполнению плана.

9. АНАЛИЗ РЕГРЕССИИ И КОРРЕЛЯЦИИ

РЕГРЕССИОННО-КОРРЕЛЯЦИОННЫЙ АНАЛИЗ

Понятие о

регрессионном

и корреляционном

анализе и его этапы

Метод аналитических группировок не дает аналитического выражения связи в виде той или иной математической формулы, функции, характеризующей и выражающей, хотя бы приближенно, механизм взаимодействия количественного признака (y) и признака-фактора (или нескольких факторных признаков — x_i): $\hat{y}_x = f(x)$. Подобного рода аналитическое выражение связи можно получить при использовании метода анализа регрессий и корреляций — регрессионно-корреляционного анализа (РКА). Он является продолжением и углублением метода аналитических группировок. РКА основан на тесном, неразрывном установлении качественного, теоретического анализа сущности изучаемых явлений с использованием приемов и методов математической статистики.

РКА заключается в построении и анализе экономико-математической модели в виде уравнения регрессии (уравнения корреля-

ционной связи), т. е. в виде той или иной функции, приближенно выражающей зависимость среднего значения результативного признака от одного или нескольких признаков-факторов: $y_{12} = f(x, z, \dots)$. Это может быть, например, линейная функция: $y_{12} = a_0 + a_1x_1 + a_2x_2$ (здесь оба факторных признака обозначены одной и той же буквой x с соответствующими номерами). Так для колхозов Восточно-Латвийской равнины Латвийской ССР О. П. Крастинем было построено следующее уравнение регрессии, выражающее зависимость средней урожайности зерновых в 1960 г. (y — центнеров с 1 га) от качества пашни (x_1 — баллов) и от количества удобрений всех видов, внесенных на 1 га (x_2 — в пересчете на действующие вещества, центнеров): $y_{12} = 0,97 + 0,26x_1 + 3,5x_2$. Из уравнения видно, что между качеством пашни и количеством удобрений, с одной стороны, и урожайностью зерновых, с другой стороны, существует прямая связь: при прочих равных условиях увеличением качества пашни на 1 балл урожайность повышается в среднем на 0,26 ц, а увеличение количества внесенных на 1 ц удобрений на 1 ц приводило к росту урожайности в среднем почти на 3,7 ц.

РКА состоит из следующих этапов (стадий): I — *предварительный (априорный) анализ*; II — *сбор информации и ее первичная обработка*; III — *построение модели (уравнения регрессии)*; IV — *оценка и анализ модели*. Этапы связаны между собой и их границы нередко переплетаются и несут условный характер.

Предварительный (априорный) анализ

При предварительном анализе прежде всего нужно в самом общем виде сформулировать задачу исследования. Например, изучение зависимости урожайности от наиболее важных факторов, оказывающих влияние на ее уровень и динамику, и т. п. В процессе априорного анализа эта задача уточняется и конкретизируется. Далее необходимо обосновать *методику измерения* результативного признака, т. е. выбрать показатель (измеритель), наиболее точно характеризующий этот признак с точки зрения поставленной задачи. Так, если результативным признаком является урожайность, то нужно решить, использовать ли для расчета валовой сбор в первоначально оприходованном поле уборки весе или же в весе после доработки, т. е. за вычетом отходов, и т. п.

Далее необходимо наметить *перечень наиболее важных и существенных факторов*, которые теоретически (или по мнению специалистов) должны влиять на результативный признак. Нужно также выбрать измерители этих факторов, т. е. факторные признаки (показатели). *Отбор факторов и их измерителей* является наиболее важной и ответственной проблемой РКА. Обычно этот отбор является многостадийным. Сначала, при априорном анализе, в перечень включаются все факторы, которые предположительно должны влиять на результативный признак. При этом в перечень могут быть включены разные варианты измерения отдельных фак-

торов (например, фактор «уровень применения удобрений» может быть измерен либо несколькими факторными признаками, каждый из которых выражает количество удобрений определенного вида — органических, минеральных и т. п., либо одним синтетическим, суммарным показателем — количеством удобрений всех видов в пересчете на действующие вещества). Если качественный (теоретический, логический) анализ не позволяет отобрать из этого предельного перечня наиболее важные факторы и такие их измерители, которые лучше отвечают задаче исследования, то на последующих этапах РКА отбор факторов и их измерителей производится на основе *сочетания качественного анализа с применением показателей и критериев математической статистики* (они будут рассмотрены ниже).

Прежде чем приступить к сбору данных, *второй и первичная обработка информации* нужно установить и уточнить *границы исследуемой совокупности* во времени и пространстве, а также *единицу* этой совокупности. Так, при изучении урожайности в качестве единицы совокупности могут быть использованы участок, бригада, колхоз или совхоз, район и т. д. В каждом конкретном случае выбор единицы должен отвечать целям и задачам анализа и возможностям получения необходимых сведений применительно к принятым нормам.

Исходная информация представляет собой характеристику каждой единицы исследуемой совокупности величиной результативного признака и одного или нескольких признаков-факторов. Исследуемая совокупность должна удовлетворять определенным условиям и требованиям. РКА основан на использовании метода средних величин как обычных, так и особых, как это будет показано далее. Поэтому *исследуемая совокупность*, во-первых, *должна быть качественно однородной*. Так, при изучении урожайности совокупность должна быть однородной в отношении категории хозяйства (колхозы или совхозы), по почвенно-климатическим условиям, по производственному направлению и специализации и т. п. При изучении производительности труда в промышленности необходима однородность предприятий в отношении специализации и концентрирования, применяемой технологии и т. п. Основным методом обеспечения однородности является группировка.

Во-вторых, *совокупность должна быть достаточно большой по числу (числу единиц или наблюдений)*, чтобы в результате действия закона больших чисел показатели регрессии и корреляции были достаточно надежными и устойчивыми и отражали объективные закономерности взаимосвязи, свободные от воздействия случайных обстоятельств.

В-третьих, *наблюдения должны быть стохастически независимыми*. Это значит, что результаты каждого наблюдения, т. е. значения признаков у той или иной единицы совокупности, не должны зависеть от результатов других наблюдений, т. е. значений данного признака у других единиц совокупности. Наконец, РКА

предполагает, что каждому значению признака-фактора (x_i) соответствует нормальное или близкое к нему распределение результативного признака (y) с одинаковой дисперсией.

Первичная обработка исходной информации заключается прежде всего в проверке ее достоверности и проверке того, соблюдаются ли в необходимых пределах условия, которым она должна удовлетворять (однородность, независимость, нормальность и др.).

В социально-экономических исследованиях обычно не все указанные условия соблюдаются достаточно строго, особенно условие нормальности и равенства дисперсий при разных значениях факторных признаков. Это накладывает определенные ограничения на применение РКА в тех случаях, когда то или иное условие резко нарушается. Однако незначительные отклонения от этих условий не являются препятствием для применения РКА, так как погрешности в результатах будут невелики.

В некоторых случаях обеспечение соблюдения одного условия может привести к нарушению другого условия. Так, если совокупность оказалась неоднородной, то ее разбивка на однородные группы может привести к тому, что число единиц в некоторых из них окажется слишком малым. Для увеличения числа наблюдений в таких случаях иногда используется так называемый метод завола-лет. Он состоит в том, что в качестве отдельных наблюдений применяются показатели по каждой единице совокупности (например, заводу, заводу и т. п.) за каждый отдельный год. Однако при этом может быть нарушено условие независимости наблюдений: если значение какого-либо признака, например уровня производительности труда, на одном заводе не зависит обычно от его значений на других заводах, то значение этого признака на данном заводе в данном году, конечно, зависит от его значений на этом заводе в предыдущие годы. Поэтому при использовании метода завола-лет применение РКА оказывается: оно требует обязательной проверки независимости наблюдений, а при наличии зависимости — специальных методов ее устранения или уменьшения влияния.

Проверка достоверности и соблюдения условий, которым должна удовлетворять исходная информация, производится путем качественного ее анализа и с помощью методов и критериев математической статистики (F -критерий Фишера, t -критерий Стьюдента и др.).

В процессе первичной обработки информации продолжается также отбор факторов и их измерителей. Включаемые в уравнение регрессии факторы должны оказывать достаточно существенное влияние на результативный признак, т. е. связь последнего с каждым факторным признаком должна быть достаточно тесной. Для предварительной проверки этого могут быть использованы рассмотренные выше методы — сопоставление параллельных рядов и особенно аналитическая группировка — простая и комбинационная.

С другой стороны, факторы, включаемые в уравнение регрессии, не должны находиться между собой в линейной функциональной или очень тесной корреляционной связи. Тесно связанные между собой факторы дублируют друг друга и искажают результаты РКА, вызывая неустойчивость коэффициентов уравнения регрессии. Поэтому нужно измерять тесноту связи каждого фактора с каждым из остальных, вычислив парные линейные коэффициенты корреляции (см. о них ниже), и исключить один или несколько тесно связанных с другим (другими) факторов.

Построение модели уравнения регрессии

При построении регрессионно-корреляционной модели (уравнения регрессии) прежде всего возникает вопрос о *типе аналитической функции*, характеризующей механизм взаимосвязи между результативным признаком и одним или несколькими признаками-факторами. Так, применительно к одному факторному признаку речь идет о том, будет ли геометрическим выражением этой аналитической функции прямая линия ($\hat{y}_x = a_0 + a_1x$), парабола второго порядка ($\hat{y}_x = a_0 + a_1x + a_2x^2$), двучленная гипербола ($\hat{y}_x = a + \frac{a_1}{x}$) или какая-либо иная кривая. Иными словами, возникает вопрос о выборе (установлении) формы связи. Выбор той или иной формы связи означает выдвижение и принятие некоторой теоретически обоснованной или практически приемлемой рабочей гипотезы о механизме взаимодействия изучаемых признаков. Поэтому этот этап РКА является очень важным и ответственным.

После того как выбрана форма связи и уравнение регрессии построено в общем виде, нужно на основе исходной информации найти числовые значения параметров уравнения регрессии (a_0, a_1, \dots, a_n), т. е. решить модель.

Выбор формы связи при построении многофакторных моделей (уравнений парной регрессии)

Выбор формы связи должен быть основан на качественном, теоретическом и логическом анализе природы изучаемых явлений, их социально-экономической сущности. Поэтому прежде всего нужно попытаться установить форму связи путем экономического и логического анализа. При построении многофакторных моделей (уравнений парной регрессии) такой анализ в ряде случаев позволяет вывести уравнение регрессии, т. е. сделать вывод формулы, приближенно характеризующей зависимость результативного признака от признака-фактора. Если бы зависимость y от x была функциональной, то теоретический анализ дал бы возможность представить y как строгую и точную функцию от x . Однако и при корреляционной связи можно нередко представить зависимость y от x аналитически в виде приближенной формулы, если абстрагироваться от того обстоятельства, что кроме данного фактора на y влияют и многие другие факторы. Это абстрагирование приведет, конечно, к некоторому упрощению действительного механизма взаимосвязи, однако оно позволит выделить и сконцентрировать внимание на том, что нас в данном случае интересует, — на зависимости y от x при прочих равных условиях.

Рассмотрим, например, вопрос о форме связи между дневной выработкой рабочего на формовке бетона (y — куб. м) и уровнем механизации труда (x — доля времени, отработанного в течение смены на механизированной формовке). Доля времени, отработанного в течение смены вручную, выразится тогда как $(1 - x)$. Обозначим дневную выработку рабочего при выполнении

работы механизированным способом через y_m , а при формовке вручную — через y_p . Тогда уровень дневной выработки бетонщика проработавшего часть смены (x) на механизированной формовке, а часть $(1-x)$ — на ручной, может быть выражен как среднее арифметическая взвешенная из дневной выработки при механизированной работе (y_m) и работе вручную (y_p) с весами x и $(1-x)$: $y = y_m x + y_p (1-x) = y_m x + y_p - y_p x = y_p + (y_m - y_p)x$.

Считая y_p и y_m величинами постоянными (средними за смену), мы видим, что зависимость дневной выработки бетонщика от уровня механизации его труда может быть выражена уравнением прямой, не проходящей через начало координат: $y = A_0 + A_1 x$, где $A_0 = y_p$, $A_1 = y_m - y_p$.

Так как $y_m > y_p$, то $A_1 > 0$, а уравнение означает, что с возранием уровня механизации труда (x) дневная выработка бетонщика увеличивается равномерно (по прямой линии).

Естественно, что у каждого бетонщика могут быть различные уровни выработки как на механизированной (y_m), так и на ручной формовке (y_p), что связано с факторами, не рассматриваемыми в данном примере, но влияющими на производительность труда (квалификация, производственный стаж и др.), вследствие чего и различные коэффициенты A_0 и A_1 найденного уравнения могут быть различными для отдельных рабочих. Поэтому при одном и том же уровне механизации труда уровень дневной выработки у отдельных бетонщиков принимает различное значение. Чтобы поставить дневную выработку бетонщиков (y) в зависимость только от уровня механизации труда (x) и исключить влияние прочих факторов, необходимо определить типичские и значения коэффициентов полученного уравнения, в данном случае средний для всех бетонщиков уровень дневной выработки на механизированной и на ручной формовке бетона, а также среднее значение разницы между уровнем дневной выработки при выполнении работы механизированным и ручным способом. Обозначим A_0 через a_0 и A_1 через a_1 . Тогда получится следующее уравнение линейной регрессии:

$$\hat{y}_x = a_0 + a_1 x. \quad (8.6)$$

Необходимо иметь в виду, что уравнение регрессии правильно выражает форму связи исследуемых показателей при условии независимости коэффициентов a_0 и a_1 от факторного признака (x) либо такой их незначительной зависимости, которой можно пренебречь.

Уравнение (8.6) похоже и непохоже на уравнения функциональной зависимости. Оно похоже на них своей формой, но отличается по своему смыслу. Объясняется это тем, что уравнение регрессии с осредненными коэффициентами отражает соотношение признаков только в общем итоге по изучаемой совокупности. Для каждой же отдельной единицы это соотношение принимает индивидуальный характер. Следовательно, уравнение регрессии показывает типичное в данных конкретных условиях соотношение между размерами факторного и результативного признаков, а не

каждого отдельного случая.

Уравнения регрессии отличаются от функциональных и в том отношении, что типические соотношения, выявляемые ими, меняются в зависимости от объема совокупности. По мере возрастания объема изучаемой совокупности повышается типичность показателей регрессии и корреляции. В этом проявляется действие закона больших чисел.

В ряде случаев теоретический анализ приводит к выводу критической формы уравнения регрессии. Рассмотрим, например, зависимость себестоимости единицы продукции от объема производства этой продукции. Себестоимость единицы продукции рассчитывается путем деления общей суммы затрат на производство данной продукции на ее количество, объем. Поэтому общая сумма затрат на производство (общая себестоимость всей продукции) делится на произведение себестоимости единицы продукции (обозначим ее y) на объем продукции (x), т. е. yx . В то же время затраты на производство можно условно и приблизительно подразделить на две части: 1) затраты, которые возрастают более или менее пропорционально увеличению объема произведенной продукции, — условно-переменные расходы (затраты на сырье и материалы, на топливо и электроэнергию для технологических целей, оплата труда основных производственных рабочих и т. п.); 2) затраты, либо совершенно не зависящие от объема производства, либо зависящие от него в незначительной степени, — условно-постоянные расходы (оплата труда инженерно-технических работников и служащих, расходы на содержание зданий и сооружений и другие административно-управленческие и общепроизводственные расходы).

Обозначим переменные расходы в расчете на единицу продукции через A_0 , тогда их общая сумма составит $A_0 x$. Общую сумму условно-постоянных расходов обозначим A_1 . Тогда общая себестоимость продукции составит: $yx = A_0 x + A_1$, откуда себестоимость единицы продукции будет равна: $y = A_0 + \frac{A_1}{x}$. Для каждого предприятия, выпускающего данную продукцию, расходы A_0 и A_1 имеют различную величину. Приняв их типичные (средние) значения ($a_0 = \bar{A}_0$ и $a_1 = \bar{A}_1$), получим следующее уравнение гиперболической регрессии (уравнение двучленной гиперболы):

$$\hat{y}_x = a_0 + \frac{a_1}{x}. \quad (8.7)$$

Поскольку $a_1 > 0$, то зависимость себестоимости единицы продукции от объема ее выпуска является обратной: с увеличением x уравнения себестоимость снижается. Однако это снижение не является равномерным, так как по мере увеличения x снижение y_x постепенно замедляется.

В случае гиперболической регрессии вида (8.7) средняя величина результативного признака при равномерном изменении x из-

меняется неравномерно, асимптотически приближаясь к a_0 . Однако в некоторых случаях теоретический и логический анализ показывает, что неравномерное изменение результативного признака должно иметь иной характер. Так, при недостаточном количестве осадков урожайность будет, естественно, очень низкая, а по мере увеличения их количества урожайность будет повышаться. Однако это повышение не будет беспредельным, так как для каждой культуры в данных конкретных условиях есть какое-то оптимальное количество осадков, при котором достигается наиболее высокая урожайность. По мере того как количество осадков будет приближаться к оптимальной величине, рост урожайности будет постепенно замедляться и прекратится совсем по достижении этого оптимума. Дальнейшее увеличение количества осадков может привести к тому, что они окажутся излишними и вредными, в результате чего урожайность будет снижаться. Такого рода зависимость можно приближенно выразить уравнением параболы второго порядка ($\hat{y}_x = a_0 + a_1x + a_2x^2$). Аналогичный характер связи можно ожидать и в ряде других случаев, например для зависимости уровня производительности труда рабочего от его возраста. С увеличением возраста и стажа работы накапливается опыт работы, совершенствуются навыки, повышается квалификация и производительность труда. Но для каждого вида работы есть, очевидно, некоторый определенный возраст, при котором выработка работника за единицу времени будет наибольшей. В дальнейшем же благоприятное влияние производственного опыта и навыков будет перекрываться неблагоприятным влиянием возрастных особенностей сложного человека, в результате чего выработка будет снижаться. Следовательно, если рассматривать зависимость производительности труда от возраста в достаточно большом диапазоне их изменения, то можно предполагать, что по мере увеличения возраста выработка рабочего будет изменяться неравномерно: сначала она будет быстро возрастать, затем этот рост будет постепенно замедляться и прекратится совсем. После этого прямая зависимость может смениться обратной, т. е. изменить свое направление: увеличение возраста будет сопровождаться некоторым снижением выработки.

В некоторых случаях теоретический анализ дает основание предполагать, что относительное изменение результативного признака должно быть пропорционально относительному приросту признака-фактора, например увеличение x на 1% должно приводить к увеличению или уменьшению y на a процентов. В таких случаях в качестве формы связи можно принять степенную функцию вида $\hat{y}_x = a_0x^{a_1}$.

Социально-экономические явления очень сложны. Как правило, мы не имеем о них исчерпывающей информации, а внутренняя логика их связей мало изучена. Факторы, влияющие на то или иное явление, взаимно переплетаются и взаимодействуют друг с другом. Поэтому очень часто не удается сделать теоретически обоснован-

ный вывод уравнения регрессии, т. е. формы связи, внутренне присущей изучаемому явлению. В ряде случаев на основе теоретического анализа могут быть высказаны лишь более или менее обобщенные предположения о том, следует ли ожидать линейную или какую-либо нелинейную (криволинейную) связь, имеет ли ожидаемая криволинейная функция экстремальные значения, асимптоты и т. п. Более того, если явление мало изучено, иногда могут быть высказаны и различные гипотезы о механизме и форме взаимодействия фактора и против каждой из которых можно привести в равной степени убедительные доводы и соображения.

Для проверки тех или иных предположений и гипотез может быть использован графический метод — построение графика группировки средних, полученных в процессе аналитической группировки (см. рис. 8.1 и 8.2). Ломаная линия, изображающая изменение групповых средних результативного признака в зависимости от изменения группировочного признака-фактора, называется эмпирической линией регрессии (эмпирической регрессией).

Форма эмпирической регрессии дает возможность проверить, существует ли фактическое соотношение признаков тому или иному теоретически предполагаемому их соотношению. При этом можно, однако, иметь в виду, что при относительно небольшом числе единиц совокупности (числе наблюдений) форма эмпирической линии регрессии может изменяться при изменении числа групп и их границ. Поэтому при небольшом числе наблюдений анализ слишком полагаться на форму эмпирической регрессии, графический метод в таких случаях может оказаться недостаточно надежным.

Если относительно формы связи могут быть выдвинуты различные теоретические гипотезы, а по виду эмпирической регрессии трудно судить о том, какой из этих гипотез наиболее соответствует фактические данные, то в этих случаях строятся и решаются уравнения регрессии с различными формами связи, а затем с помощью специальных статистико-математических критериев оценивается их адекватность и выбирается та форма связи, которая обеспечивает наилучшую аппроксимацию (приближение) и достаточную статистическую достоверность и надежность.

Выбрав тем или иным путем форму связи и построив уравнение регрессии в общем виде, необходимо далее найти числовые значения его параметров, т. е. решить модель. Напомним, что вследствие влияния прочих факторов параметры уравнения для отдельных единиц совокупности могут быть различными, так что задача заключается в том, чтобы найти их типичные средние для всей совокупности значения. При этом в отличие от обычного порядка нахождения средних величин значения рассматриваемого варьирующего признака (в данном случае индивидуальных значения параметров у отдельных единиц совокупности) здесь не известны. Кроме того, процесс осреднения параметров

уравнения есть в то же время процесс осреднения значений результирующего признака, причем эта средняя (\hat{y}_x) является не постоянной, а переменной величиной, меняющейся в связи с изменением факторного признака (x). Таким образом, процесс осреднения при построении уравнения регрессии носит особый характер.

Чтобы произвести такое осреднение, обычно исходят из свойства средней арифметической, которое заключается в том, что сумма квадратов отклонений вариантов от средней арифметической меньше, чем сумма квадратов отклонений от любой другой величины. Применительно к парной регрессии это означает, что минимальной должна быть $\sum (y - \hat{y}_x)^2$. Заменяя затем \hat{y}_x выражением соответствующим выбранной форме связи [например, при линейной связи $\hat{y}_x = a_0 + a_1x$ будем иметь $\sum (y - a_0 - a_1x)^2$], и рассматривая полученную сумму квадратов отклонений как функцию параметров, можно найти те их значения, при которых эта сумма будет минимальной. Это достигается путем приравнивания нулю первых частных производных по каждому параметру и решения полученных при этом системы так называемых нормальных уравнений. Такой способ расчета параметров называется методом наименьших квадратов.

Если исходные данные (значения x и y) нанести на график в виде точек в прямоугольной системе координат, то получим картину корреляции (см. рис. 8.3). Если бы зависимость y от x была функциональной, то все точки были бы расположены либо на какой-то прямой, либо на той или иной кривой. При корреляционной связи вследствие влияния прочих факторов точки не лежат на прямой или кривой, но все же их расположение обнаруживает определенную тенденцию. Так, на рис. 8.3 видно, что по мере повышения уровня механизации труда (x) дневная выработка бетонщиков (y) в общем и целом повышается. Как показал сделанный нами ранее теоретический анализ, зависимость здесь должна выражаться уравнением прямой $\hat{y}_x = a_0 + a_1x$, где параметр a_0 характеризует средний для всех бетонщиков уровень дневной выработки на ручной формовке, а параметр a_1 — превышение средней дневной выработки на механизированной формовке над средней выработкой при работе вручную (при условии, что уровень механизации труда (x) выражен в виде доли).

Расчет параметров прямой методом наименьших квадратов дает такие их значения, при которых прямая на графике пройдет наиболее близко к точкам, изображающим исходные фактические данные. Наибольшая близость в данном случае означает, что если в уравнение прямой последовательно подставлять фактические значения x у каждой единицы совокупности и рассчитывать соответствующие значения результирующего признака (\hat{y}_x), то сумма квадратов отклонений фактических значений y от расчетных (теоретических), т. е. $\sum (y - \hat{y}_x)^2$, будет при этом меньше, чем для любой другой прямой, которую можно провести на корреляционном поле. Аналогично обстоит дело и при других (криволинейных) фор-

мах связи: расчет параметров кривой методом наименьших квадратов дает возможность найти ту кривую, которая по сравнению с другими кривыми данного вида проходит наиболее близко к фактическим корреляционным полям, изображающим фактические значения признаков, т. е. дает наименьшую сумму квадратов отклонений $\sum (y - \hat{y}_x)^2$.

Расчет параметров прямой линейной регрессии по сгруппированным данным

Рассмотрим вывод нормальных уравнений для линейной парной регрессии $\hat{y}_x = a_0 + a_1x$. Сумма квадратов отклонений y от \hat{y}_x составляет $S = \sum (y - \hat{y}_x)^2 = \sum (y - a_0 - a_1x)^2$. Первая частная производная этой суммы по a_0 равна: $S_{a_0} = 2\sum (y - a_0 - a_1x) \cdot (-1) = 2\sum (a_0 + a_1x - y)$. Приравняв ее нулю и разделив обе части равенства на 2, получим $\sum (a_0 + a_1x - y) = 0$, откуда после раскрытия скобок будем иметь $\sum a_0 + a_1\sum x - \sum y = 0$. Заменяя сложение одинаковых слагаемых a_0 произведением a_0n и вынеся a_1 за знак суммы, получим первое уравнение системы: $a_0n + a_1\sum x = \sum y$. Первая частная производная S по a_1 равна: $S_{a_1} = 2\sum (y - a_0 - a_1x) \cdot (-x) = 2\sum (a_0x + a_1x^2 - yx)$. После приравнивания ее нулю, деления на 2, раскрытия скобок и вынесения a_0 и a_1 за знак суммы получим второе уравнение системы: $a_0\sum x + a_1\sum x^2 = \sum yx$. Таким образом, система нормальных уравнений для парной линейной регрессии такова:

$$\begin{cases} a_0n + a_1\sum x = \sum y, \\ a_0\sum x + a_1\sum x^2 = \sum yx. \end{cases}$$

Решение этой системы в общем виде дает следующие значения параметров:

$$a_0 = \frac{\sum y \cdot \sum x^2 - \sum yx \cdot \sum x}{n \sum x^2 - \sum x \cdot \sum x}; \quad (8.8a)$$

$$a_1 = \frac{n \sum yx - \sum y \cdot \sum x}{n \sum x^2 - \sum x \cdot \sum x}. \quad (8.8б)$$

Поскольку вторые частные производные S по a_0 и a_1 положительны ($S_{a_0} = 2\sum x^2$), функция S при данных значениях a_0 и a_1 имеет минимум, а не максимум.

Рассчитаем параметры линейного уравнения регрессии, выражающего зависимость дневной выработки рабочего на формовке бетона (y — куб. м) от уровня механизации труда (x_1 — процент) по исходным данным табл. 8.4. Для подсчета необходимых сумм ($\sum x_1$, $\sum y$, $\sum yx_1$, $\sum x_1^2$) составим расчетную таблицу (значения x_1^2 , x_2 и другие, приводимые в последних графах, будут использованы в дальнейших расчетах).

Таблица 4

РАСЧЕТ СУММ ДЛЯ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ПАРНОГО ЛИНЕЙНОГО
УРАВНЕНИЯ РЕГРЕССИИ ПО НЕСГРУППИРОВАННЫМ ДАННЫМ

№ п/п	x_i	y_i	x_i^2	yx_i	y^2	x_i	x_i^2	yx_i
1	35	3,0	1 225	105,0	9,00	2	4	6,0
2	59	6,5	3 481	383,5	42,25	3	9	19,5
3	44	4,8	1 936	211,2	23,04	3	9	14,4
...
40	67	5,5	4 489	368,5	30,25	2	4	11,0
50	57	5,9	3 249	336,3	34,81	3	9	17,7
Итого	2 800	250,0	171 536	14 752,0	1 337,96	150	476	781,4

Используя найденные суммы, по формуле (8.8) найдем параметры уравнения линейной регрессии:

$$a_0 = \frac{250 \times 171\,536 - 14\,752 \times 2\,800}{50 \times 171\,536 - 2\,800 \times 2\,800} = \frac{1\,578\,400}{736\,800} = 2,142;$$

$$a_1 = \frac{50 \times 14\,752 - 250 \times 2\,800}{50 \times 171\,536 - 2\,800 \times 2\,800} = \frac{37\,600}{736\,800} = 0,051.$$

Следовательно, уравнение регрессии таково:

$$\hat{y}_{x_i} = 2,142 + 0,051x_i.$$

Графическое изображение корреляционного поля и найденного уравнения регрессии показано на рис. 8.3.

При выводе уравнения регрессии, характеризующего зависимость дневной выработки от уровня механизации труда (рис. 8.3), было показано, что

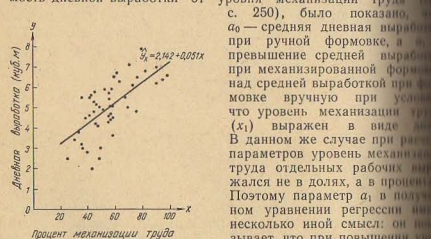


Рис. 8.3. Зависимость дневной выработки от уровня механизации труда (корреляционное поле и уравнение регрессии)

с. 250), было показано, что a_0 — средняя дневная выработка при ручной формовке, а a_1 — превышение средней выработки при механизированной формовке над средней выработкой при ручной формовке при условии, что уровень механизации труда (x_i) выражен в виде доли. В данном же случае при расчете параметров уровень механизации труда отдельных рабочих выражался не в долях, а в процентах. Поэтому параметр a_1 в полученном уравнении регрессии имеет несколько иной смысл: он показывает, что при повышении уровня механизации на 1% дневная выработка рабочего возрастает в среднем на 0,051 м³. Если бы уровень

параметров x_i был выражен в долях, то параметр a_0 остался бы прежним, а параметр a_1 был бы в 100 раз больше (1 м³) и показывал бы прирост средней дневной выработки при повышении уровня механизации с 0 до 1, т. е. при переходе от ручной к полностью механизированной работе. Поскольку в данном рассматриваемом примере (табл. 8.4) минимальный уровень механизации (25%) довольно далек от нуля, т. е. нет рабочих, которые работали вручную в течение всего рабочего дня, интерпретация параметров уравнения регрессии (при измерении x_i в долях) такова. В течение той части рабочего дня, когда рабочие работали вручную, выработка одного рабочего в расчете на полный рабочий день составляла в среднем 2,14 м³ (a_0), а в течение той части дня, когда рабочие были заняты механизированной работой, выработка (тоже в расчете на полный рабочий день одного рабочего) была в среднем на 5,10 м³ выше (a_1), т. е. составляла 7,24 м³.

Нужно иметь в виду, что в тех случаях, когда уравнение линейной регрессии получено не путем вывода, как это было сделано в данном случае, а на основе графика групповых средних или путем эмпирического подбора формы связи, параметр a_0 часто не имеет содержательного статистико-экономического смысла и носит чисто расчетный характер. Интерес представляет тогда только параметр a_1 , показывающий, как в среднем изменяется y при изменении x на единицу.

Построим, например, уравнение регрессии, характеризующее зависимость дневной выработки от квалификации (тарифного разряда) рабочего. Очевидно, что повышение квалификации рабочего должно приводить к росту производительности труда. Однако, требуемая чисто теоретически, трудно сказать, будет ли рост средней дневной выработки при повышении разряда на единицу пропорционально равномерно, по прямой линии, или же неравномерно, т. е., например, прирост выработки при переходе от 3-го разряда к 4-му будет не равен ее приросту при повышении разряда со 2-го до 3-го. Возможно, например, что рабочие наиболее высокой квалификации значительно старше рабочих низкой и средней квалификации, так что при повышении разряда рост выработки будет происходить сначала быстро, а затем более медленно вследствие неблагоприятного влияния возрастных особенностей пожилого человека.

Следовательно, теоретически можно предполагать как линейную, так и криволинейную — параболическую зависимость. Многие данные здесь от конкретных условий исследуемой совокупности, от взаимосвязи данного фактора (разряда) с другими факторами.

Приним в качестве объекта исследования совокупность 50 рабочих-бетонщиков (см. табл. 8.4) и построив график групповых средних, характеризующих изменение средней дневной выработки в группах рабочих разной квалификации (см. табл. 8.7 и рис. 8.2), мы видим, что эмпирическая линия регрессии является почти прямой (более точные значения групповых средних равны 3,76; 5,03 и 6,18); прирост выработки при повышении разряда со 2-го до 3-го

Таблица 8.13

ТАБЛИЦА ФОРМ ДЛЯ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ЛИНЕЙНОЙ РЕГРЕССИИ
ПО ДАННЫМ КОРРЕЛЯЦИОННОЙ ТАБЛИЦЫ

число рабочих по уровню квалификации (м²)	2-4	4-6	6-8				
число рабочих по уровню квалификации (м²)	середина интервала			Итого m_x	$x m_x$	$x^2 m_x$	Итого $\sum x m_y$
x	y						
34-45	35	105 7 735	175 7 1225	—	14	490	17 150
46-55	55	165 3 495	275 17 4675	385 3 1155	23	1 265	69 575
56 и более	82	—	410 4 1 640	574 9 5 166	13	1 066	87 412
Итого	m_y $\sum m_y$	10 30	28 140	12 84	50 254	2 821	174 137

В нашем примере

$$a_0 = \frac{254 \times 174 137 - 15 091 \times 2 821}{50 \times 174 137 - 2 821 \times 2 821} = 2,216;$$

$$a_1 = \frac{50 \times 15 091 - 254 \times 2 821}{50 \times 174 137 - 2 821 \times 2 821} = 0,051.$$

Расчеты значительно упрощаются при использовании способа моментов (особенно при равных интервалах). Так, в нашем примере, при равных интервалах для x расчет будет таким (вместо x и y берутся x' и y' ; расположение показателей в клетках таблицы прежнем) (см. табл. 8.14).

Используя формулу (8.9), где вместо x и y берутся x' и y' , найдем условные параметры:

$$a_0' = \frac{2 \times 30 - 16 \times (-12)}{50 \times 30 - (-12) \times (-12)} = 0,186;$$

$$a_1' = \frac{50 \times 16 - 2 \times (-12)}{50 \times 30 - (-12) \times (-12)} = 0,608.$$

Следовательно,

$$\hat{y}_x = 0,186 + 0,608x', \text{ или } \frac{\hat{y}_x - 5}{2} = 0,186 + 0,608 \cdot \frac{x - 62,5}{25},$$

откуда получим

$$\hat{y}_x = 2,332 + 0,049x.$$

составляет 1,27 м³, а с 3-го до 4-го — 1,15 м³. Учитывая, что увеличение прироста очень незначительно и что в узком диапазоне изменения признака-фактора (2—4) параболу с небольшой кривизной вполне можно заменить прямой, в порядке первого приближения примем для уравнения регрессии линейную форму связи $\hat{y}_x = a_0 + a_1 x_2$, где x_2 — тарифный разряд.

Используя итоги табл. 8.12, по формуле (8.8) вычислим параметры a_0 и a_1 и построим уравнение регрессии: $\hat{y}_x = 1,377 + 1,208x_2$. Параметр a_0 имеет здесь чисто расчетное значение, так как равный нулю разряда нет и быть не может ($a_0 = \hat{y}_x$ при $x_2 = 0$). Параметр же a_1 показывает, что при повышении квалификации на разряд дневная выработка увеличивается в среднем на 1,208 м³.

Расчет параметров парной линейной регрессии по сгруппированным данным

Использование первичных *несгруппированных* данных дает наиболее точную величину параметров уравнения регрессии. Однако при большом числе единиц совокупности (наблюдений) расчет параметров по первичным данным становится очень трудным и громоздким. Значительное упрощение достигается путем *комбинационной группировки* по факторному и результативному признакам, однако точность расчета при этом снижается, так как вместо индивидуальных значений признаков используются *середины интервалов*. Лишь при достаточно большом числе единиц и их равномерном распределении внутри групп значения параметров будут мало отличаться от тех, которые были получены по несгруппированным данным.

Результаты комбинационной группировки оформляются в виде *корреляционной таблицы (решетки)*, по данным которой и производится расчет параметров уравнения регрессии. Сначала сделаем такой расчет без использования способа моментов.

В середине каждой клетки, образованной пересечением строк и графы, показаны частоты m_{yx} (числа рабочих), в верхнем левом углу — произведения yx , а в правом нижнем углу — произведение $xy m_{yx}$ (так, в 1-й клетке 1-й строки число рабочих $m_{yx} = 7$, $yx = 3 \times 35 = 105$; $xy m_{yx} = 735$). Сумма этих произведений показана в последней графе (так, по 1-й строке $735 + 1225 = 1960$). Значения $\sum y^2 m_y$ нужны в дальнейшем для измерения тесноты связи.

Параметры уравнения регрессии получим по формулам:

$$a_0 = \frac{\sum y m_y \cdot \sum x^2 m_x - \sum xy m_{yx} \cdot \sum x m_x}{n \sum x^2 m_x - \sum x m_x \cdot \sum x m_x}; \quad (8.9a)$$

$$a_1 = \frac{n \sum xy m_{yx} - \sum y m_y \cdot \sum x m_x}{n \sum x^2 m_x - \sum x m_x \cdot \sum x m_x}, \quad (8.9b)$$

где $n = \sum m_x = \sum m_y$.

Таблица 1

РАСЧЕТ СУММ ДЛЯ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ СПОСОБОМ МОМЕНТОВ

Уровень механизации труда (процентов)		Уровень дневной выработки (м ³)			Итого m_x	$x^2 m_x$	$(x^2)^2 m_x$	
		2-4	4-6	6-8				
x	y	3	5	7	m_x	$x^2 m_x$	$(x^2)^2 m_x$	
	$y' = \frac{y-5}{2}$							
	$x' = \frac{x-62,5}{25}$	-1	0	1				
25-50	37,5	-1	1 9 9 0	0 12 0 0	—	21	-21	21
50-75	62,5	0	1 0 0	0 14 5 0	20	0	0	0
75 и более	87,5	1	—	0 2 0	1 7 7	9	9	9
Итого		m_y $y' m_y$ $(y')^2 m_y$	10 -10 10	28 0 12	12 12 22	50	-12	30

Отличие параметров этого уравнения от параметров предыдущего уравнения объясняется тем, что для обеспечения равных интервалов изменены границы групп.

Выше было показано, что зависимость себестоимости единицы продукции от объема ее производства может быть приближенно выражена уравнением гиперболической регрессии вида $\hat{y}_x = a + \frac{a_1}{x}$. Оно отличается

$$a_0 = \frac{\sum y \cdot \sum \frac{1}{x^2} - \sum \frac{y}{x} \cdot \sum \frac{1}{x}}{n \sum \frac{1}{x^2} - \sum \frac{1}{x} \cdot \sum \frac{1}{x}}, \quad (8.10a)$$

$$a_1 = \frac{n \sum \frac{y}{x} - \sum y \cdot \sum \frac{1}{x}}{n \sum \frac{1}{x^2} - \sum \frac{1}{x} \cdot \sum \frac{1}{x}}. \quad (8.100)$$

Например, по 20 заводам имеются данные о выпуске продукции и себестоимости ее единицы, изображенные на рис. 8.4 в виде

На графике построена также эмпирическая регрессия, форма которой подтверждает, что исходные данные не противоречат предположению о гиперболической форме связи.

Для расчета нужных сумм построим таблицу, в которую включим также необходимые в дальнейшем значения y^2 .

Таблица 8.15

СЧЕТ СУММ ДЛЯ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ГИПЕРБОЛЫ

№ п/п	Объем произ- водства (тыс. единиц) x	Себестоимость единицы (руб.) y	$\frac{1}{x}$	$\frac{1}{x^2}$	$\frac{y}{x}$	y^2
1	5,0	30	0,200	0,0400	6,00	900
2	5,5	28	0,182	0,0331	5,09	784
3	5,6	29	0,179	0,0319	5,18	841
4	10,0	23	0,100	0,0100	2,30	529
5	11,0	24	0,091	0,0083	2,18	576
Итого		515	2,748	0,3939	71,71	13341

Подставив в формулу (8.10) вычисленные суммы, найдем, что $-17,76$ и $a_1 = 58,13$, откуда искомое уравнение регрессии таково: $-17,76 + \frac{58,13}{x}$. На рис. 8.4 показан график этого уравнения.

Если бы в расчетах объем выпуска (x) выражался в единицах, а не в тысячах единиц, параметр a_0 остался бы таким же, а пара-

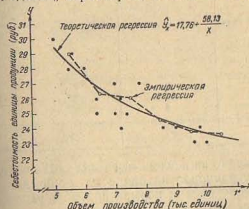


Рис. 8.4. Зависимость себестоимости единицы продукции от объема производства

был бы в 1000 раз больше. Следовательно, в среднем по заводам затраты, пропорциональные объему продукции, составляют 17,76 руб. на каждую единицу продукции, а общая сумма постоянных расходов равна 58 130 руб. Чем больше объем

продукции, тем меньшая часть этой суммы падает на каждую единицу продукции, в результате чего зависимость себестоимости единицы продукции от ее объема и оказывается обратной.

Если форма связи выражается параболой второго порядка $\hat{y}_x = a_0 + a_1x + a_2x^2$, то для определения ее параметров, удовлетворяющих требованиям метода наименьших квадратов, нужно решить следующую систему нормальных уравнений (при негруппированных данных):

$$\left. \begin{aligned} a_0n + a_1 \sum x + a_2 \sum x^2 &= \sum y, \\ a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3 &= \sum xy, \\ a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4 &= \sum x^2y. \end{aligned} \right\} \quad (8.11a, б, в)$$

Допустим, по 30 колхозам, находящимся в примерно одинаковых почвенно-климатических условиях, имеются данные об общем количестве внесенных на 1 га удобрений (в пересчете на действующие вещества) и урожайности зерновых. Эти данные изображены на рис. 8.5 в виде точек корреляционного поля.

Как видно из графика, по мере увеличения количества внесенных на каждый гектар удобрений урожайность имеет тенденцию к повышению. Эта тенденция подтверждается и изменением групповых средних.

ЗАВИСИМОСТЬ УРОЖАЙНОСТИ ОТ КОЛИЧЕСТВА УДОБРЕНИЙ

Внесено удобрений на 1 га (ц)	Число колхозов	Средняя урожайность (ц с 1 га)	Увеличение урожайности по сравнению с предыдущей группой (ц)
Менее 0,65	6	18,0	—
0,65—0,95	6	22,8	4,8
0,95—1,25	10	26,7	3,9
1,25—1,55	8	29,8	3,1
Итого	30	25,0	—

Однако, как видно из последней графы таблицы, прибавка урожайности не пропорциональна приросту количества удобрений: при относительно невысоких нормах внесения удобрений эта прибавка выше, а при высоких нормах — ниже.

Теоретический и логический анализ также приводит к выводу о том, что по мере увеличения норм внесения удобрений их эффект не будет оставаться постоянным, так как в каждом конкретном случае существует та или иная оптимальная норма удобрений, обеспечивающая максимальную урожайность. Следовательно, в порядке первого приближения можно считать, что уравнение регрессии в нашем случае будет представлять ветвь параболы второго порядка.

Обозначив количество удобрений через x , а урожайность через y , определим параметры параболы по исходным данным нашего графика, для чего сначала найдем нужные суммы.

Таблица 8.17

СУММ ДЛЯ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ПАРАБОЛЫ ВТОРОГО ПОРЯДКА

Число колхозов	x	y	x^2	x^3	x^4	xy	x^2y	\hat{y}_x
6	0,4	14	0,16	0,064	0,0256	5,6	2,24	14,98
6	0,5	16	0,25	0,125	0,0625	8,0	4,00	17,11
10	0,5	19	0,25	0,125	0,0625	9,5	4,75	17,11
8	1,4	32	1,96	2,744	3,8416	44,8	62,72	30,02
8	1,5	30	2,25	3,375	5,0625	45,0	67,50	30,77
30	30,0	750	32,90	38,484	47,0762	791,1	899,95	749,95

Подставив найденные суммы в систему нормальных уравнений (8.11), получим:

$$\left. \begin{aligned} 30a_0 + 30a_1 + 32,9a_2 &= 750, \\ 30a_0 + 32,9a_1 + 38,484a_2 &= 791,1, \\ 32,9a_0 + 38,484a_1 + 47,0762a_2 &= 899,95. \end{aligned} \right\}$$

Решение этой системы уравнений дает следующие значения параметров: $a_0 = 5,086$; $a_1 = 27,511$; $a_2 = -6,927$. Следовательно, уравнение регрессии имеет такой вид:

$$\hat{y}_x = 5,086 + 27,511x - 6,927x^2.$$

Подставляя в это уравнение соответствующие значения x , найдем теоретические (ожидаемые, расчетные) значения \hat{y}_x . Так, при $x = 0,4$ получим $\hat{y}_x = 5,086 + 27,511 \times 0,4 - 6,927 \times (0,4)^2 = 14,98$; при $x = 0,5$ — соответственно 17,11 и т. д. (см. последнюю графу табл. 8.17). График уравнения регрессии показан на рис. 8.5.

Значения \hat{y}_x , полученные по уравнению регрессии, показывают, какую в среднем величину урожайности можно ожидать при данном количестве удобрений и прочих равных (средних для данной совокупности) условиях. Из графика и приведенных значений \hat{y}_x видно, что в условиях исследуемой совокупности повышение урожайности за счет увеличения количества удобрений происходит и не равномерно: равные изменения факторного признака x вызывают неравные, меняющиеся для разных уровней x изменения \hat{y}_x . Так, при увеличении x с 0,4 до 0,5 (см. табл. 8.17) ожидаемая прибавка урожайности составляет 2,13 ц (17,11—14,98), при увеличении x с 0,5 до 0,6 — 1,99 ц, а при увеличении x с 1,4 до 1,5 — только 0,77 ц (30,77—30,02). Скорость изменения урожайности является

здесь переменной, она постепенно замедляется при увеличении x . Размер же этого замедления — величина постоянная: каждое новое увеличение x на 0,1 дает прибавку урожайности, которая 0,14 и меньше предыдущей $[2a_2 \times (0,1)^2]$.

Для определения параметров уравнения регрессии, выраженного степенной функцией вида $\hat{y}_x = a_0 x^{a_1}$, нужно привести эту функцию к линейному виду путем логарифмирования: $\lg \hat{y}_x = \lg a_0 + a_1 \lg x$. Полученное уравнение отличается от уравнения обычной линейной регрессии тем, что вместо y , x и a_0 содержит их логарифмы. Поэтому формулы параметров степенной регрессии можно получить, заменив в формуле (8.8) показатели y , x и a_0 их логарифмами (величина a_1 логарифмом не заменяется):

$$\lg a_0 = \frac{\sum \lg y \cdot \sum (\lg x)^2 - \sum \lg y \cdot \lg x \cdot \sum \lg x}{n \sum (\lg x)^2 - \sum \lg x \cdot \sum \lg x}, \quad (8.13a)$$

$$a_1 = \frac{n \sum \lg y \cdot \lg x - \sum \lg y \cdot \sum \lg x}{n \sum (\lg x)^2 - \sum \lg x \cdot \sum \lg x}. \quad (8.13b)$$

Параметр a_1 здесь показывает, на сколько процентов изменяется в среднем y при изменении x на один процент.

Оценка и анализ однофакторных моделей (уравнений парной регрессии). Измерение тесноты связи

После построения модели (уравнения регрессии) ее необходимо оценить и проверить. Прежде всего нужно проверить, согласуются ли знаки параметров с теоретическими представлениями и соображениями о направлении влияния признака-фактора на результативный признак (показатель). Иными словами, нужно проверить, как (в каком направлении) изменяется \hat{y}_x при изменении x , соответствует ли направление этого изменения тому направлению, которое предполагалось при теоретическом и логическом анализе.

Для однофакторных моделей (уравнений парной регрессии) такое соответствие чаще всего имеет место. Так, в уравнении линейной регрессии $\hat{y}_x = 2,142 + 0,051x_1$, характеризующем зависимость дневной выработки рабочего-бетонщика (\hat{y}_x) от уровня механизации его труда (x_1), параметр $a_1 > 0$, т. е. при повышении уровня механизации выработка, как и ожидалось, также увеличивается. В другом уравнении линейной регрессии ($\hat{y}_x = 1,377 + 1,208x_2$), выражающем зависимость дневной выработки от квалификации рабочего (тарифного разряда — x_2), параметр a_1 также положительный, что вполне согласуется с экономическими соображениями о том, что повышение квалификации рабочего должно приводить к росту производительности труда.

В уравнении регрессии ($\hat{y}_x = 17,76 + \frac{58,13}{x}$), характеризующем зависимость себестоимости единицы продукции (\hat{y}_x) от объема ее

производства (x), с увеличением x величина $\frac{58,13}{x}$ уменьшается и

также уменьшается, так как $a_1 = +58,13$. Это тоже полностью соответствует теоретическим представлениям об обратной зависимости себестоимости единицы продукции от объема ее производства.

При построении уравнения параболической регрессии, характеризующего зависимость урожайности (\hat{y}_x) от количества внесенных на 1 га удобрений (x), предполагалось, что с увеличением x урожайность должна повышаться, но ее увеличение будет постепенно замедляться, так что уравнение регрессии будет изображаться частью той ветви параболы второго порядка, которая направлена вниз (к оси абсцисс ОХ). Как видно из рис. 8.5, знаки параметров a_1 и a_2 оказались таковы ($a_1 > 0$; $a_2 < 0$), что изменение \hat{y}_x соответствует тому, которое предполагалось.

Иногда может оказаться, что знаки параметров уравнения регрессии не соответствуют теоретическим соображениям о направлении изменения \hat{y}_x . Например, вместо ожидавшегося положительного параметра a_1 он оказался отрицательным, или наоборот. В подобных случаях нужно попытаться выяснить причины выявившегося несоответствия. Ими могут быть либо недоброкачественность исходной информации (ошибки в первичных данных, недостаточное число наблюдений, неоднородность совокупности и т. п.), либо ошибки в расчетах, либо, наконец, неучтенные при теоретическом анализе взаимосвязи рассматриваемых признаков с другими. Последнее обстоятельство может привести к тому, что априорные предположения о направлении или характере изменения \hat{y}_x в конкретных условиях исследуемой совокупности не подтвердятся, ранее принята гипотеза окажется недостаточной обоснованной или ошибочной.

Важное место при оценке модели занимает измерение тесноты связи. Как и при методе аналитических группировок, для измерения тесноты связи используются показатели вариации результативного признака (y).

Общая дисперсия $\sigma_y^2 = \frac{\sum (y - \bar{y})^2}{n}$ является мерой колеблемости фактических (эмпирических) значений признака y около их средней величины. Она характеризует общую вариацию результативного признака y , объясняемую влиянием всех факторов, от которых он зависит. Отклонения $y - \bar{y}$ объясняются тем, что сочетание значений факторов, влияющих на y , у каждой единицы совокупности является индивидуальным, различным.

Кроме общей дисперсии, на основе уравнения регрессии и отклонений $\hat{y}_x - y$ можно вычислить средний квадрат этих отклонений, т. е. средний квадрат отклонений теоретических значений \hat{y}_x , рассчитанных по уравнению регрессии, от их средней величины (средняя величина \hat{y}_x равна общей средней y):

$$\sigma_{\hat{y}_x}^2 = \frac{\sum (\hat{y}_x - y)^2}{n}. \quad (8.13)$$

Это — мера колеблемости теоретических значений признака около их средней величины.

На основе уравнения регрессии и отклонений $y - \hat{y}_x$ можно также вычислить средний квадрат этих отклонений, т. е. средний квадрат отклонений фактических значений результативного признака от теоретических его значений, полученных путем подстановки в уравнение регрессии соответствующих значений признака фактора:

$$\sigma_y^2 = \frac{\sum (y - \hat{y}_x)^2}{n}. \quad (8.10)$$

Это — мера колеблемости фактических значений y около соответствующих теоретических значений, т. е. около линии регрессии. В математической статистике доказывается, что

$$\sigma_y^2 = \sigma_{\hat{y}_x}^2 + \sigma_{\text{ост}}^2. \quad (8.11)$$

Рассмотрим смысл дисперсий $\sigma_{\hat{y}_x}^2$ и $\sigma_{\text{ост}}^2$ на примере зависимости урожайности от количества удобрений (см. рис. 8.5).

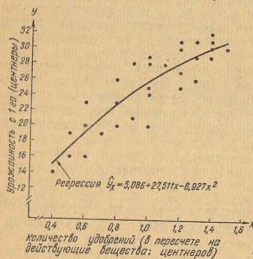


Рис. 8.5. Зависимость урожайности от количества удобрений, внесенных на 1 га

Если бы урожайность зависела только от количества удобрений, т. е. если бы связь между этими признаками была не корреляционной, а функциональной и совершенно точно выражалась бы уравнением построенной параболы, то каждому значению x (количество удобрений) соответствовало бы вполне определенная величина y (урожайность). Все фактические значения y совпадали бы с теоретическими (расчетными): $y = \hat{y}_x$, т. е. все точки на рис. 8.5 были бы расположены на параболе (линии регрессии). Отклонения $y - \hat{y}_x$ были бы тогда равны нулю, дисперсия $\sigma_{\text{ост}}^2$ тоже равнялась бы нулю, отклонения $\hat{y}_x - \bar{y}$ совпадали бы с отклонениями $y - \bar{y}$, а дисперсия $\sigma_{\hat{y}_x}^2$ с общей дисперсией σ_y^2 . Следовательно, равенство $\sigma_{\hat{y}_x}^2 = \sigma_y^2$ свидетельствует о том, что связь y с x функциональная и вся вариация y объясняется фактором x .

Рассмотрим теперь противоположный случай. Если бы урожайность совершенно не зависела от количества удобрений, а число наблюдений было достаточно велико, то при любом количестве удобрений (при любом значении x) урожайность (y) была бы одинаковой или примерно одинаковой, равной или близкой к средней величине (\bar{y}). Иначе говоря, при отсутствии связи и достаточно большом числе наблюдений все значения \hat{y}_x , независимо от величины x , были бы равны \bar{y} , т. е. $\hat{y}_x = \bar{y}$. Уравнение регрессии выражалось бы тогда прямой, параллельной оси абсцисс, а отклонения $y - \hat{y}_x$ и дисперсия $\sigma_{\hat{y}_x}^2$ были бы равны нулю. Отклонения же $y - \bar{y}$

совпадали бы с отклонениями $y - \bar{y}$, а дисперсия $\sigma_{\text{ост}}^2$ с общей дисперсией σ_y^2 . Следовательно, равенства $\sigma_{\hat{y}_x}^2 = 0$ и $\sigma_{\text{ост}}^2 = \sigma_y^2$ свидетельствуют о том, что y не связан с x , что вся вариация y объясняется другими, не входящими в уравнение регрессии факторами.

Фактически в нашем примере с урожайностью и удобрениями, как видно из рис. 8.5, мы имеем дело не с рассмотренными крайними случаями (функциональная связь и полное отсутствие связи), а с некоторым промежуточным случаем, когда связь есть, но она является не функциональной, а корреляционной, т. е. кроме количества удобрений (x) на урожайность влияют и другие факторы. Поскольку y зависит от x , регрессия не совпадает с прямой $y = \bar{y}$, поскольку же кроме x на y влияют и другие факторы, фактические значения урожайности не совпадают со значениями, вычисленными по уравнению регрессии, и точки на графике не лежат на линии регрессии.

Следовательно, то обстоятельство, что линия регрессии не совпадает с прямой $y = \bar{y}$, а отклонения $\hat{y}_x - \bar{y}$ не равны нулю, объясняется зависимостью y от x , т. е. тем, что вариация, изменение x сопровождается изменением, вариацией \hat{y}_x . Иными словами, вариация теоретических значений \hat{y}_x их колеблемость около \bar{y} , измеряемая дисперсией $\sigma_{\hat{y}_x}^2$, объясняется влиянием на y только признака фактора x , входящего в уравнение регрессии. Поэтому дисперсия $\sigma_{\hat{y}_x}^2$, которая характеризует вариацию y , объясняемую только фактором x , называется факторной дисперсией.

С другой стороны, если форма связи выбрана правильно (т. е. зависимость урожайности от количества удобрений при прочих равных условиях действительно выражается параболой второго порядка), отклонения фактических значений y от теоретических \hat{y}_x объясняются тем, что кроме x на y влияют и другие факторы, вариация которых сопровождается колебаниями y около \hat{y}_x , т. е. около линии регрессии. Иными словами, колеблемость y около \hat{y}_x , измеряемая дисперсией $\sigma_{\text{ост}}^2$, объясняется влиянием на y остальных факторов, от которых зависит результативный признак, т. е. прочих факторов, не входящих в уравнение регрессии. Из равенства (8.15)

следует, что эта дисперсия может быть получена как остаток от вычитания факторной дисперсии из общей: $\sigma_y^2 = \sigma_y^2 - \sigma_{\hat{y}_x}^2$. Показатель дисперсия σ_y^2 называется *остаточной дисперсией*. Она характеризует остаточную вариацию y , т. е. вариацию, не объясняемую фактором x , включенным в уравнение регрессии.

Для измерения тесноты связи между y и x логично поэтому использовать отношение факторной дисперсии к общей дисперсии резуль-
тативного признака. Это отношение называется *теоретическим индексом детерминации* (i^2):

$$i^2 = \frac{\sigma_{\hat{y}_x}^2}{\sigma_y^2} \quad (8.16)$$

Из смысла дисперсий σ_y^2 и $\sigma_{\hat{y}_x}^2$ вытекает, что *теоретический индекс детерминации показывает, какая часть общей вариации резуль-
тативного признака y объясняется признаком-фактором x , включенным в соответствующее уравнение регрессии.*

Если форма связи, принятая в уравнении регрессии, не соответствует действительному механизму взаимосвязи, то остаточная дисперсия будет выражать не только вариацию y , объясняемую прочими (кроме x) факторами, но и ее характером. Поэтому и факторная дисперсия также не измерит точно вариацию, объясняемую фактором x . С другой стороны, если фактор x коррелирован с y , то с изменением x будут изменяться и эти связанные с ним факторы, так что факторная дисперсия будет характеризовать вариацию y , объясненную не только x , но и связанными с ним факторами. Рассматривая пока лишь значение регрессии и корреляции, мы абстрагируемся от этого обстоятельства. Уточнение смысла индекса детерминации будет сделано при рассмотрении множественной регрессии и корреляции.

Если связь между y и x отсутствует и при любой величине x теоретические значения $\hat{y}_x = \bar{y}$, отклонения $\hat{y}_x - \bar{y}$ и факторная дисперсия будут равны нулю. Поэтому и теоретический индекс детерминации также будет равен нулю. Линия же регрессии совпадает с прямой $\hat{y}_x = \bar{y}$, параллельной оси абсцисс.

При функциональной связи между признаками фактические значения y точно равны теоретическим \hat{y}_x , так что отклонения $\hat{y}_x - \bar{y}$ совпадают с отклонениями $y - \bar{y}$, а факторная дисперсия $\sigma_{\hat{y}_x}^2$ — с общей дисперсией σ_y^2 . Поэтому индекс детерминации будет равен единице. Все точки на графике в этом случае лежат на линии ре-
грессии.

Если же связь между признаками есть, но она является не функциональной, а корреляционной, то по мере ее усиления, т. е. по мере повышения тесноты связи, индекс детерминации увеличивается, приближаясь к единице, а по мере ослабления — уменьшается, приближаясь к нулю. Таким образом, индекс детермина-

ция измеряя тесноту связи, характеризует степень близости корреляционной связи к строгой функциональной.

В качестве показателя тесноты связи используется также квадратный корень из индекса детерминации, называемый *индексом корреляции* (i), или *теоретическим корреляционным отношением*. Индекс корреляции также может принимать значения от 0 до 1.

Индекс детерминации (i^2) и индекс корреляции (i) могут быть использованы для измерения тесноты связи при любой ее форме — как прямолинейной (линейной), так и криволинейной.

Если, однако, y нелинейно зависит от параметров уравнения регрессии (например, если уравнение регрессии выражается степенной функцией $\hat{y}_x = a_0 x^{a_1}$, т. е. зависимость y от x нелинейна), равенство 8.15 может не соблюдаться, так как средняя величина \hat{y}_x не равна \bar{y} . В таких случаях для приближенного измерения тесноты связи, как и при построении уравнения регрессии, функция приводится к линейному виду (например, путем логарифмирования $\lg \hat{y}_x = a_0 \lg x + a_1$, $\lg x$) и при расчете i^2 в формулах (8.13) и (8.16) вместо y , \hat{y}_x и \bar{y} используются их логарифмы.

Определим тесноту связи между дневной выработкой рабочего (y) и уровнем механизации труда (x_1), зависимость между которыми выражается линейной регрессией $\hat{y}_x = 2,142 + 0,051x_1$ (см. табл. 8.4). Подставляя в это уравнение фактические значения x_1 (см. табл. 8.4), найдем теоретические значения \hat{y}_x . Расчет суммы квадратов отклонений, необходимой для вычисления факторной дисперсии, показан в табл. 8.18 ($y = 250 : 50 = 5$).

Таблица 8.18

Расчет $\sum (\hat{y}_x - \bar{y})^2$ для регрессии $\hat{y}_x = 2,142 + 0,051x_1$

№ п/п	x_1	\hat{y}_x	$\hat{y}_x - \bar{y}$	$(\hat{y}_x - \bar{y})^2$
1	35	3,93	-1,07	1,1449
2	59	5,15	0,15	0,0225
3	44	4,39	-0,61	0,3721
...
49	67	5,56	0,56	0,3136
50	57	5,05	0,05	0,0025
Итого	2800	249,92	—	38,2976

Факторная дисперсия $\sigma_{\hat{y}_x}^2 = \frac{38,2976}{50} = 0,7660$. Общая дисперсия $\sigma_y^2 = \overline{y^2} - (\bar{y})^2 = 1337,96 : 50 - 5^2 = 1,7592$ ($\sum y^2 = 1337,96$; см. табл. 8.12).

Отсюда индекс детерминации равен:

$$r^2 = \frac{\sum_{\bar{y}}^2}{\sum_{\bar{y}}^2} = \frac{0,7660}{1,7592} = 0,435.$$

Следовательно, 43,5% общей вариации выработки (y) объясняется вариацией уровня механизации труда (x_1).

Расчет факторной дисперсии по формуле (8.13), требующий вычисления всех теоретических значений \bar{y}_x , является очень трудным, что заставляет прибегать к округлениям. Более удобно для расчета факторной дисперсии использовать другую формулу, не требующую вычисления \bar{y}_x . Для линейной регрессии эта формула такова:

$$\sigma_{\bar{y}_x - a_0 + a_1 x}^2 = (a_0 \sum y + a_1 \sum yx) : n - (\sum y : n)^2. \quad (8.17)$$

В нашем примере имеем ($\sum y = 250$; $\sum yx = 14\,752$, см. табл. 8.19):
 $\sigma_{\bar{y}_x}^2 = (2,142 \times 250 + 0,051 \times 14\,752) : 50 - (250 : 50)^2 = 0,7570$.

Расхождение этого результата с предыдущим объясняется округлениями при расчетах.

Для гиперболической регрессии формула факторной дисперсии имеет такой вид:

$$\sigma_{\bar{y}_x - a_0 + \frac{a_1}{x}}^2 = (a_0 \sum y + a_1 \sum \frac{y}{x}) : n - (\sum y : n)^2. \quad (8.18)$$

Так, для построенного выше уравнения регрессии, характеризующего зависимость себестоимости единицы продукции от объема ее производства ($\bar{y}_x = 17,76 + \frac{58,13}{x}$), факторная дисперсия равна (необходимые суммы взяты из табл. 8.15):

$$\sigma_{\bar{y}_x}^2 = (17,76 \times 515 + 58,13 \times 71,71) : 20 - (515 : 20)^2 = 2,6826.$$

Общая дисперсия $\sigma_y^2 = 13\,341 : 20 - (515 : 20)^2 = 3,9875$. Следовательно, индекс детерминации равен: $r^2 = 2,6826 : 3,9875 = 0,673$, т. е. более 67% общей вариации себестоимости объясняется объемом производства продукции.

Для параболической регрессии второго порядка факторная дисперсия может быть исчислена по такой формуле:

$$\sigma_{\bar{y}_x - a_0 + a_1 x + a_2 x^2}^2 = (a_0 \sum y + a_1 \sum yx + a_2 \sum yx^2) : n - (\sum y : n)^2. \quad (8.19)$$

Для регрессии, характеризующей зависимость урожайности от количества удобрений ($\bar{y}_x = 5,086 + 27,511x - 6,927x^2$), факторная дисперсия составляет (суммы взяты из табл. 8.17): $\sigma_{\bar{y}_x}^2 = (5,086 \times 750 + 27,511 \times 791,1 - 6,927 \times 899,95) : 30 - (750 : 30)^2 = 19,8160$.

Общая дисперсия урожайности, исчисленная по первичным данным, здесь равна 25,6; следовательно, индекс детерминации равен: $r^2 = 10,8166 : 25,6 = 0,774$, т. е. более 77% общей вариации урожайности объясняется колеблемостью количества удобрений, внесенных на 1 га.

При линейной форме связи показатели ее тесноты (индекс детерминации и индекс корреляции) называются коэффициентом детерминации (r^2) и коэффициентом корреляции (r). Коэффициент детерминации имеет тот же смысл, что и индекс детерминации, но используется только в случае линейной связи. Обычно сначала рассчитывается коэффициент корреляции. Он может быть исчислен по одной из следующих формул:

$$r = \frac{\sum (y - \bar{y})(x - \bar{x})}{n \sigma_y \sigma_x} = \frac{\sum yx - \sum y \cdot \sum x}{\sqrt{[n \sum y^2 - (\sum y)^2] \cdot [n \sum x^2 - (\sum x)^2]}}. \quad (8.20a, б, в)$$

Коэффициент корреляции может принимать значения от -1 до $+1$, включая и 0 . Отрицательные значения указывают на обратную связь, положительные — на прямую связь. При $r = 0$ линейная связь отсутствует. При $r = \pm 1$ связь является функциональной. По абсолютной величине коэффициент корреляции равен индексу корреляции, если связь линейна.

Найдем коэффициент корреляции между дневной выработкой рабочего-бетонщика и уровнем механизации труда (x_1), используя для расчета по последней формуле итоговые данные табл. 8.12:

$$r = \frac{50 \times 14\,752 - 250 \times 2\,800}{\sqrt{(50 \times 1337,96 - 250^2) \times (50 \times 171\,536 - 2\,800^2)}} = 0,6605.$$

Коэффициент детерминации $r^2 = 0,436$. Этот результат почти совпадает с исчисленным выше индексом детерминации (расхождение вызвано округлениями при расчетах).

По данным табл. 8.12 можно найти и коэффициент корреляции между дневной выработкой и квалификацией (тарифным разрядом — x_2):

$$r = \frac{50 \times 781,4 - 250 \times 150}{\sqrt{(50 \times 1337,96 - 250^2) \times (50 \times 476 - 150^2)}} = 0,6566,$$

откуда $r^2 = 0,431$, т. е. 43% вариации выработки (y) объясняется вариацией тарифных разрядов.

Коэффициент корреляции можно рассчитать и по данным корреляционной таблицы (решетки):

$$r = \frac{n \sum x_1 y_{12} - \sum x_1 y_1 \cdot \sum y_{12}}{\sqrt{[n \sum y_{12}^2 - (\sum y_{12})^2] \cdot [n \sum x_1^2 - (\sum x_1)^2]}}. \quad (8.21)$$

Эту формулу можно применять как в том случае, когда в качестве x и y взяты середины интервалов, так и при использовании

способа моментов. В последнем случае вместо x и y в формуле берутся x' и y' .

Так, например, по данным табл. 8.14 коэффициент корреляции будет равен:

$$r = \frac{50 \times 16 - 2 \times (-12)}{\sqrt{(50 \times 22 - 2^2)(50 \times 30 - (-12)^2)}} = 0,6759.$$

Этот результат менее точен, чем тот, который был получен ранее на основе первичных негруппированных данных.

При анализе и экономической интерпретации уравнений регрессии используются также коэффициенты эластичности. Результативного признака относительно признака-фактора. Коэффициент эластичности показывает, на сколько процентов изменится в среднем y при изменении x на 1%. Коэффициент эластичности ($\hat{\epsilon}$) вычисляется по такой формуле:

$$\hat{\epsilon} = \frac{\partial \hat{y}_x}{\partial x} \cdot \frac{x}{\hat{y}_x}, \quad (8.15)$$

где $\frac{\partial \hat{y}_x}{\partial x}$ — первая производная \hat{y}_x по x .

При большинстве форм связи коэффициент эластичности является переменным, т. е. изменяется с изменением признака-фактора x . Так, для линейной регрессии $\hat{\epsilon} = a_1 x$ ($a_0 + a_1 x$). Применительно к уравнению, выражающему зависимость выработки от уровня механизации труда получим: $\hat{\epsilon} = 0,051 x_1$; $(2,142 + 0,051 x_1)$. При $x_1 = 50\%$, $\hat{\epsilon} = 0,54$, а при $x_1 = 56\%$, $\hat{\epsilon} = 0,57$. Следовательно, если уровень механизации труда повышается с 50 до 50,5% (т. е. на 1,0% раза, или на 1% прежней величины), то выработка повышается в среднем на 0,54% прежнего уровня.

При степенной форме связи ($\hat{y}_x = a_0 x^{a_1}$) коэффициент эластичности является постоянной величиной, равной a_1 .

Являясь относительными величинами, коэффициенты эластичности позволяют сравнивать процентное изменение результативного признака при изменении различных факторов на 1%. Так, если при повышении среднего уровня механизации с 56 до 56,56% ($56 \times 1,01$) выработка повышается на 0,57%, то при 0,72%. Однако при таких сопоставлениях нужно учитывать также относительную колеблемость факторов в изучаемой совокупности, т. е. их коэффициенты вариации. В данном случае коэффициент вариации уровня механизации равен 30,7%, а коэффициент вариации тарифных разрядов — только 24,0%. Следовательно, увеличение на 1% одного фактора относительно незначительно такому же увеличению другого.

Показатели регрессии и корреляции — параметры уравнения регрессии, индексы или коэффициенты детерминации и корреляции, — исчисленные для ограниченной по объему совокупности, могут быть искажены действием случайных факторов. Поэтому нужно проверить, насколько эти показатели характерны для того комплекса условий, в которых находится исследуемая совокупность, не являются ли они результатом стечения случайных обстоятельств.

Проверка значимости (существенности) показателя регрессии и корреляции производится с помощью критерия математической статистики: критерия t Стьюдента, дисперсионного критерия F Фишера и др.

При линейной зависимости для оценки значимости параметров уравнения регрессии используется критерий t , который для a_0 рассчитывается по формуле

$$t = \frac{\frac{\hat{a}_0}{\sigma_{\hat{a}_0}}}{\sigma_{\hat{a}_0}}, \text{ а для } a_1 - \text{ по формуле } t_1 = \frac{\hat{a}_1}{\sigma_{\hat{a}_1}} \cdot \sigma_x. \text{ Вычислен-}$$

ные по этим формулам значения t сравниваются затем с критическими их значениями при принятом уровне значимости (существенности) α и числе степеней свободы $k = n - 2$. Критические значения t находятся по таблице распределения t -студента. В социально-экономических исследованиях уровень значимости α обычно принимается равным 0,05. Если расчетное значение t больше критического, параметр признается значимым (отклоняется гипотеза о том, что параметр в действительности равен нулю и лишь в силу случайных обстоятельств оказался равным проверяемой величине).

Для уравнения линейной регрессии $\hat{y}_x = 2,142 + 0,051 x_1$, характеризующего зависимость дневной выработки от уровня механизации труда (см. с. 269), $t = \frac{2,142}{0,2142} \cdot \sqrt{50 - 2} = 0,997 = 14,88$ (остаточная дисперсия $\sigma_y^2 = \sigma_y^2 - \sigma_{\hat{y}_x}^2 = 17992 - 0,7660 = 0,9932$ (см. с. 269), откуда $\sigma_y = 0,997$); $t_1 = \frac{0,051}{0,002} \cdot \sqrt{50 - 2} = 17,167$; $0,997 = 6,08$ [$\sigma_x^2 = \Sigma x_1^2 : n - (\Sigma x_1 : n)^2 = 171\,536 : 50 - (2800 : 50)^2 = 294,72$ (см. с. 256), откуда $\sigma_x = 17,167$].

По таблице распределения Стьюдента для $k = 50 - 2 = 48$ и $\alpha = 0,05$ найдем критическое значение $t = 2,01$. Следовательно, оба параметра значимы.

Для каждого параметра можно построить доверительные интервалы (границы) $a_i \pm t_{\alpha/2} \sigma_{\hat{a}_i}$, где t — коэффициент доверия по распределению Стьюдента при $k = n - 2$ степенях свободы; μ_j — средняя ошибка a_j , для a_0 равная $\mu_0 = \frac{\sigma_{\hat{a}_0}}{\sqrt{n - 2}}$, для $a_1 - \mu_1 = \sigma_{\hat{a}_1} \cdot \sigma_x \sqrt{n - 2}$.

При линейной зависимости для оценки значимости коэффициента корреляции (коэффициента детерминации) также можно использовать критерий t , который в этом случае рассчитывается по формуле

$$t = \frac{r \sqrt{n - 2}}{\sqrt{1 - r^2}}.$$

Так, для коэффициента корреляции между дневной выработкой и уровнем механизации, исчисленного по негруппированным данным (см. с. 271), получим $t = \frac{0,6605 \sqrt{50 - 2}}{\sqrt{1 - 0,4363}} = 6,10$. Это значительно больше критического значения t для 48 степеней свободы и $\alpha = 0,05$ (2,01), что свидетельствует о значимости коэффициента корреляции и существенности связи между выработкой и уровнем механизации.

При криволинейной связи для оценки значимости индекса корреляции (коэффициента детерминации) используется критерий F , который может быть вычислен по одной из следующих формул: $F = \frac{\frac{\sigma_{\hat{y}_x}^2}{\sigma_y^2} \cdot \frac{n - m}{m - 1}}{\frac{n - m}{1 - F}}$, $F = \frac{n - m}{1 - F}$.

В числитель параметров в уравнении регрессии. Расчетное значение F сравнивается с критическим (табличным) для принятого уровня значимости α и числе степеней свободы $k_1 = m - 1$ и $k_2 = n - m$.

Для зависимости себестоимости единицы продукции от объема ее производства ($n = 20$, $r^2 = 0,673$; $m = 2$) получим $F = \frac{0,673}{1 - 0,673} \times \frac{20 - 2}{2 - 1} = 37,05$. Критическое значение F при $\alpha = 0,05$ ($k_1 = 1$; $k_2 = 18$) равно 4,41. Связь себестоимости с объемом производства продукции существенна.

Аналогично производится проверка значимости эмпирического коэффициента детерминации (и эмпирического корреляционного отношения). При этом $k_1 = m - 1$; $k_2 = n - m$, где m — число групп по признаку-фактору x .

Показатели тесноты связи (индексы и коэффициенты корреляции и детерминации) используются не только для оценки анализа уже построенной модели (уравнения регрессии), но и в процессе построения модели — для выбора оптимального ее варианта, если теоретический анализ не позволяет отдать предпочтение ни одному из возможных вариантов формы связи. Так, например, если ни теоретический анализ, ни эмпирическая линия регрессии не позволяют дать однозначный ответ о возможной форме связи, можно построить уравнения регрессии с различными формами связи и оценить их сравнительную пригодность на основе сопоставления индексов корреляции или детерминации. При этом, если число наблюдений недостаточно велико, величины этих показателей необходимо скорректировать, учитывая степень свободы. Практически более удобно корректировать величины остаточных дисперсий, для чего их нужно умножить на n и разделить на $n - m$ (m — число параметров в уравнении регрессии).

Выше для зависимости урожайности от количества удобрений было построено уравнение параболической регрессии, причем остаточная дисперсия при $n = 30$ составляла $\sigma_a^2 = \sigma_y^2 - \sigma_{\hat{y}}^2 = 25,6 - 19,8166 = 5,7834$. Умножив эту величину на $n = 30$ и разделив на $n - m = 30 - 3 = 27$, получим 6,426. Если на основе тех же данных построить линейную регрессию, то $\sigma_a^2 = 6,1838$, а после умножения на 30 и деления на $n - m = 30 - 2 = 28$ скорректированная остаточная дисперсия равна 6,626. Так как скорректированная остаточная дисперсия при линейной связи больше, чем при параболической, а индекс детерминации соответственно меньше, то более предпочтительной является параболическая регрессия.

Показатели тесноты связи используются также для отбора факторов при построении многофакторных моделей (уравнений множественной регрессии) и при подготовке исходной информации для математического программирования.

Построение многофакторных моделей (уравнений множественной регрессии)

Социально-экономические явления отличаются большой сложностью. Их уровень формируется под влиянием целого комплекса переплетающихся и взаимодействующих между собой факторов, действующих с разной силой и в разных направлениях. Поэтому построение однофакторных моделей — парных уравнений регрессии — обычно оказывается недостаточным. Сложные взаимодействия данного фактора с другими могут сделать неточными, искаженными показатели парной регрессии и корреляции. Сама специфика корреляционных связей требует включения в модель наиболее важных и существенных факторов. Отбор таких

факторов и является важнейшей проблемой при построении уравнений множественной регрессии.

Как уже говорилось выше, отбор факторов производится на основе качественного, теоретического анализа с одновременным использованием статистико-математических критериев. Наиболее целесообразен трехстадийный отбор. На первой стадии — при априорном анализе — на факторы, включаемые в предварительный их перечень, можно не накладывать никаких особых ограничений (могут включаться различные варианты измерителей одного и того же фактора и т. п.). Далее, на второй стадии производится сравнительная оценка и отсев части факторов. Это делается на основе сочетания качественного анализа с анализом парных коэффициентов и индексов корреляции и оценкой их существенности (значимости). Для этого составляется матрица парных коэффициентов корреляции, измеряющих тесноту линейной связи каждого фактора с результативным признаком и с каждым из остальных признаков-факторов (корреляционная матрица):

x_0	x_1	x_2	\dots	x_j	\dots	x_n
x_0	1	r_{01}	r_{02}	\dots	r_{0j}	r_{0n}
x_1	r_{10}	1	r_{12}	\dots	r_{1j}	r_{1n}
x_2	r_{20}	r_{21}	1	\dots	r_{2j}	r_{2n}
\dots						
x_i	r_{i0}	r_{i1}	r_{i2}	\dots	r_{ij}	r_{in}
\dots						
x_n	r_{n0}	r_{n1}	r_{n2}	\dots	r_{nj}	1

Здесь x_0 — результативный признак (y); x_1, x_2, \dots, x_n — признаки-факторы; r_{ij} — парный коэффициент корреляции между x_i и x_j ($r_{11} = r_{22} = \dots = r_{nn} = 1$; $r_{01} = r_{10}$ и т. д.).

Корреляционная матрица позволяет выявить факторы, которые находятся между собой в тесной линейной корреляционной взаимосвязи, близкой к функциональной. Обычно такой взаимосвязью признается та, при которой $|r| > 0,8$ (иногда при более строгом подходе $|r| > 0,7$). При наличии таких связей между факторными признаками один или несколько из них нужно ис-

При двух факторах система уравнений (8.23) примет такой вид:

$$\left. \begin{aligned} a_0 + a_1 \sum x_1 + a_2 \sum x_2 &= \sum y, \\ a_0 \sum x_1 + a_1 \sum x_1^2 + a_2 \sum x_1 x_2 &= \sum y x_1, \\ a_0 \sum x_2 + a_1 \sum x_1 x_2 + a_2 \sum x_2^2 &= \sum y x_2. \end{aligned} \right\} \quad (8.24, 6 \text{ стр.})$$

В расчетной табл. 8.12 имеются все необходимые для этой системы суммы, подставив которые получим:

$$\left. \begin{aligned} 50a_0 + 2800a_1 + 150a_2 &= 250, \\ 2800a_0 + 171\,536a_1 + 8\,672a_2 &= 14\,752, \\ 150a_0 + 8\,672a_1 + 476a_2 &= 781,4. \end{aligned} \right\}$$

Решение этой системы дает $a_0 = 0,5002$; $a_1 = 0,0356$; $a_2 = 0,8351$. Следовательно, уравнение множественной регрессии таково:

$$\hat{y}_{12} = 0,5002 + 0,0356x_1 + 0,8351x_2.$$

Параметр a_0 здесь не имеет содержательного смысла и носит чисто расчетный характер. Параметр $a_1 = 0,0356$ показывает, что повышение уровня механизации труда на 1% при фиксированном (постоянном) значении тарифного разряда приводит к увеличению дневной выработки в среднем на 0,0356 м³. Параметр $a_2 = 0,8351$ означает, что при повышении квалификации на один тарифный разряд и при фиксированном уровне механизации труда дневная выработка увеличивается в среднем на 0,8351 м³.

Таким образом, коэффициенты при x_1 в уравнении множественной линейной регрессии показывают, насколько в среднем изменится результативный признак при увеличении соответствующего фактора на единицу и при фиксированном (постоянном) значении других факторов, входящих в уравнение регрессии.

Ранее нами были построены парные регрессии, характеризующие зависимость выработки (y) от уровня механизации ($y_{x_1} = 2,142 + 0,051x_1$) и от тарифного разряда ($\hat{y}_{x_2} = 1,377 + 1,208x_2$). Сравнивая параметры a_1 и a_2 в уравнении множественной регрессии с соответствующими коэффициентами парных регрессий, мы видим, что в множественной регрессии коэффициенты при x_1 и x_2 значительно меньше, чем коэффициенты при этих факторах в парных регрессиях (0,0356 против 0,0510 и 0,835 против 1,208). В связи с этим могут возникнуть такие вопросы. На сколько же кубометров увеличивается в среднем выработка при повышении уровня механизации на 1%: на 0,0510 м³, как это показывает парная регрессия, или только на 0,0356 м³, как это вытекает из уравнения множественной регрессии? Аналогично: на сколько увеличивается в среднем выработка за счет повышения квалификации на один разряд: на 1,208 м³ или только на 0,835 м³? Чем вызваны эти расхождения между значениями параметров в уравнениях множественной и парных регрессий?

Для ответа на эти вопросы найдем уравнение регрессии, характеризующее линейную зависимость уровня механизации труда (x_1) от тарифного разряда (x_2), т. е. $x_1 = a_0 + a_1x_2$, и измерим силу связи между этими признаками. Используя данные табл. 8.12 и приняв x_1 в качестве y и x_2 в качестве x , по формуле (8.8) найдем: $a_0 = 24,615$; $a_1 = 10,462$, откуда $\hat{x}_1 = 24,615 + 10,462x_2$. Таким же путем по формуле 8.20в получим парный коэффициент корреляции: $r = \frac{50 \times 8672 - 2800 \times 150}{\sqrt{(50 \times 171\,536 - 2800^2)(50 \times 476 - 150^2)}} = 0,4394$, откуда коэффициент детерминации $r^2 = 0,193$. Следова-

тельно, между уровнем механизации и тарифным разрядом имеется прямая линейная связь: при повышении разряда на единицу уровня механизации труда увеличивается в среднем почти на 10,5%, причем более 19% общей колеблемости уровня механизации объясняется вариацией тарифных разрядов.

Эта линейная зависимость между двумя факторами, влияющими на результативный признак — выработку (y), и является причиной того, что коэффициенты при x_1 и x_2 в уравнении множественной регрессии не совпадают с соответствующими параметрами в уравнениях парных регрессий. Коэффициенты при x_1 и x_2 в множественной регрессии показывают, насколько в среднем повышается выработка за счет увеличения соответствующего фактора на единицу при фиксированном, т. е. постоянном уровне другого фактора, входящего в уравнение регрессии. Что же касается коэффициентов при x_1 и x_2 в уравнениях парных регрессий, то при истолковании их смысла фиксирование других факторов на постоянном уровне не предполагается. Это означает, что коэффициенты парных регрессий показывают, как изменяется y при увеличении данного фактора на единицу и при одновременном соответствующем изменении корреляционно связанных с ним других факторов. Так, в нашем примере увеличение тарифного разряда на единицу приведет к повышению выработки в среднем на 1,208 м³, если при том условии, что одновременно изменится и величина других факторов, которые корреляционно связаны с разрядом (x_2). Если мы видели, что в условиях исследуемой совокупности увеличение разряда на единицу сопровождается повышением уровня механизации труда на 10,46%. Следовательно, увеличение выработки на 1,208 м³ объясняется не только повышением квалификации как таковой, но и ростом уровня механизации труда, которым сопровождается повышение квалификации (тарифного разряда).

Таким образом, коэффициент 1,208 в уравнении парной регрессии характеризует не эффект повышения квалификации самой по себе, а условный эффект этого фактора, т. е. эффект, в котором отражается как непосредственное (чистое) влияние квалификации (x_2) на выработку, так и влияние других факторов, корреляционно связанных с x_2 и через этот фактор также влияющих на выработку. Иначе обстоит дело с коэффициентами при x_1 и x_2 в уравнении множественной регрессии. Каждый из этих коэффи-

циентов тоже характеризует не только непосредственное влияние соответствующего фактора, не только его чистый эффект, но также и влияние других факторов, корреляционно связанных с данным, не входящих в уравнение регрессии. Однако влияние второго фактора, включенного в уравнение регрессии, при этом элиминировано, устранено. Следовательно, хотя параметры 0,0356 и 0,1275 в уравнении множественной регрессии и не отражают только непосредственный, чистый эффект соответствующего фактора, все же эффект, который они характеризуют «очищен» от влияния второго фактора, входящего в уравнение регрессии. Если бы в уравнении регрессии кроме x_1 и x_2 был включен еще один существенный фактор, корреляционно связанный с x_1 и x_2 , то в новом уравнении коэффициенты при x_1 и x_2 вновь изменились бы и стали еще точнее отражать непосредственное влияние факторов, их более чистой эффективность. И чем полнее в уравнении множественной регрессии представлены наиболее существенные и важные факторы, от которых зависит результативный признак, тем точнее коэффициенты регрессии выражают чистую эффективность факторов, тем менее условной становится характеристика этой эффективности, если форма связи выбрана правильно.

Для того чтобы отличать друг от друга показатели регрессии и корреляции при различном числе факторов и при различных признаках, принятых в качестве результативного (зависимого) признака, в ряде случаев целесообразно использовать более гибкую систему обозначений. При этом все признаки обозначаются одной буквой x с соответствующим подстрочным номером, причем вместо прежнего обозначения y используется x_0 . Параметры уравнения регрессии, теоретические (расчетные) значения результативного признака, а также дисперсии и показатели корреляции (коэффициенты и индексы корреляции и детерминации) имеют тройную подстрочную нумерацию. Первым указывается номер результативного (зависимого) признака, затем — номер (или номера) факторного признака, влияние (вариацию) которого характеризует (отражает) данный показатель, т. е. параметр уравнения дисперсии или индекс корреляции. После этого в скобках указываются номера других признаков-факторов, которые включены в уравнение регрессии, но влияние которых в данном случае элиминировано, устранено. В соответствии с этим уравнение регрессии характеризующее зависимость уровня механизации (x_1) от тарифного разряда (x_2), будет записано так: $\hat{x}_{12} = a_{01(2)} + a_{12}x_2$; уравнение же, выражающее зависимость выработки ($y = x_0$) от разряда, — так: $\hat{x}_{02} = a_{02(1)} + a_{02}x_2$. Наконец, полученное выше уравнение множественной регрессии будет выглядеть так:

$$\hat{x}_{012} = a_{0(12)} + a_{01(2)}x_1 + a_{02(1)}x_2$$

Используя указанные обозначения, вернемся к вопросу о соотношении параметров парных и множественной регрессий. В математической статистике делается, что при линейных связях $a_{02} = a_{02(1)} + a_{12}a_{01(2)}$. Здесь a_{02} — условный

эффект влияния x_2 на x_0 (парная регрессия), $a_{02(1)}$ и $a_{01(2)}$ — условно-чистые эффекты влияния x_2 и x_1 на x_0 (множественная регрессия), a_{12} — условный эффект влияния x_2 на x_1 (парная регрессия x_{12}). В нашем примере (см. с. 278) имеем: $0,0356 + 10,462 \cdot 0,0356 = 1,2075$ (расхождение с полученным выше значением вызвано округлениями при расчетах).

Следовательно, если абстрагироваться от прочих факторов, не включенных в уравнение множественной регрессии, то условный эффект фактора x_2 , т. е. его влияние на x_0 , складывается из двух частей: 1) $a_{02(1)}$ — условно-чистый эффект влияния x_2 на x_0 через x_1 , элиминированное влияние x_1 ; 2) a_{02} — эффект влияния x_2 , т. е. его условно прямое, непосредственное влияние на x_0 ; 2) a_{02} — дополнительное, косвенное влияние x_2 на x_0 через x_1 , объясняемое косвенно x_2 и x_1 и представляющее как бы «наслоение» на условно-чистый эффект (при отсутствии линейной связи между x_2 и x_1 параметр $a_{12} = 0$ и это «наслоение» будет отсутствовать). Схематически это можно изобразить так, как показано на рис. 8.6, где видно, что влияние x_2 на x_0 идет как бы по двум каналам: непосредственно и через x_1 , а третий канал (a_{02}) — это равнодействующая.

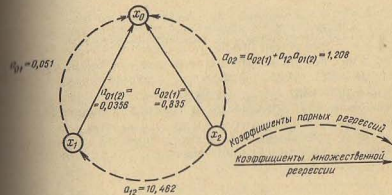


Рис. 8.6. Условный и условно-чистый эффекты влияния факторов на результативный признак x_0

Оценка и анализ влияния факторных признаков (уравнений множественной регрессии)

В условиях, когда взаимосвязи изучаемых признаков тесно переплетаются и имеют различное направление и форму, а параметры уравнения множественной регрессии характеризуют не только непосредственное влияние соответствующих факторов, но и влияние корреляционно связанных с ними других факторов, не включенных в уравнение регрессии, тогда как влияние факторов, включенных в это уравнение, элиминировано. — в этих условиях значения некоторых параметров в уравнении множественной регрессии иногда не соответствуют теоретическим и логическим представлениям и соображениям о направлении влияния соответствующих факторов. Как уже говорилось, причины такого несоответствия могут быть различными. Недостатки исходной информации (малое число наблюдений, неоднородность совокупности и т. п.) и ошибки при построении модели (при отборе факторов, при выборе формы связи) могут привести к тому, что уравнение регрессии окажется недоброкачественным: знаки пара-

метров будут искажать действительное направление влияния соответствующих факторов или другие особенности изменения результирующего признака. Однако ошибочными могут оказаться как знаки параметров, так, наоборот, и те априорные теоретические и логические соображения, с которыми сопоставляются параметры уравнения регрессии: эти априорные соображения могут оказаться недостаточно обоснованными, в них могут быть упущены и не учтены какие-либо взаимодействия факторов и тенденции, не совпадающие с предполагавшейся основной и очевидной тенденцией. Может, например, оказаться, что за факторно предполагавшееся направление влияния которых не подтвердилось в модели, скрывается действие других факторов, корреляционно связанных с ними, но не включенных в модель.

Из формулы $a_{02} = a_{02(1)} + a_{12}a_{01(2)}$ видно, что коэффициенты при одном и том же факторе в уравнении парной и множественной регрессии могут иметь различные знаки, а если $|a_{02(1)}| < |a_{12}a_{01(2)}|$, они тоже имеют разные знаки.

Как и при парной регрессии, на основе уравнения множественной регрессии общую дисперсию результирующего признака ($\sigma_y^2 = \sigma_0^2$), так как $y = x_0$ можно разложить на две части: 1) средний квадрат отклонений теоретических значений от их средних величин — факторную дисперсию $\sigma_{02...s}^2$ и 2) средний квадрат отклонений фактических значений результирующего признака от теоретических его значений — остаточную дисперсию $\sigma_a^2 = \sigma_{0(12...s)}^2$:

$$\sigma_0^2 = \sigma_{02...s}^2 + \sigma_{0(12...s)}^2, \quad (8.23)$$

$$\sigma_{02...s}^2 = \frac{\sum (\hat{x}_{02...s} - \bar{x}_0)^2}{n}, \quad (8.25a)$$

$$\sigma_{0(12...s)}^2 = \frac{\sum (x_0 - \hat{x}_{02...s})^2}{n}. \quad (8.25b)$$

Факторная дисперсия $\sigma_{02...s}^2$ характеризует вариацию результирующего признака (x_0), которая при данной форме связи объясняется факторами x_1, x_2, \dots, x_s , включенными в уравнение регрессии. Остаточная дисперсия $\sigma_{0(12...s)}^2$ характеризует остаточную вариацию x_0 , т. е. вариацию, объясняемую прочими, не включенными в уравнение регрессии факторами.

Факторная и остаточная дисперсии имеют указанный смысл лишь при условии, что связь между факторами, включенными и не включенными в уравнение регрессии, отсутствует (или мы от нее абстрагировались). При наличии такой связи факторная дисперсия, как и параметры уравнения, косвенно отражает и вариацию не включенных в модель факторов.

Для измерения тесноты связи между x_0 и факторами, включенными в уравнение регрессии, используется, как и ранее, отношение факторной дисперсии к общей дисперсии результирующего признака. Это отношение называется совокупным индексом детерминации (индексом множественной детерминации):

$$I_{02...s}^2 = \frac{\sigma_{02...s}^2}{\sigma_0^2}. \quad (8.26)$$

$I_{02...s}^2$ может принимать значения от 0 до 1 и показывает, какая часть общей вариации x_0 объясняется факторами, включенными в уравнение регрессии (при данной форме связи).

В качестве показателя тесноты связи используется также квадратный корень из $I_{02...s}^2$, называемый совокупным индексом корреляции (I). I^2 и I могут быть использованы для измерения тесноты связи при любой ее форме — как линейной, так и криволинейной.

Факторную дисперсию можно определить, не вычисляя предельно $x_{012...s}$, по рабочей формуле, аналогичной формуле (8.17). При линейной связи x_0 со всеми факторами эта формула имеет вид:

$$\sigma_{02...s}^2 = (a_0 \sum x_0 + a_1 \sum x_1 x_1 + \dots + a_s \sum x_s x_s) n - (\sum x_0 \cdot n)^2. \quad (8.27)$$

Найдем с помощью этой формулы факторную дисперсию для указанного выше уравнения множественной регрессии, используя данные табл. 8.12 и учтя, что $y = x_0$: $\sigma_{02}^2 = (0,5002 \times 250 + 0,0356 \times 14752 + 0,8351 \times 7814) : 50 - (250 : 50)^2 = 1,0554$.

Отсюда совокупный индекс детерминации равен: $I_{02}^2 = 1,0554 : 1,7592 = 0,600$ ($\sigma_0^2 = 1,7592$; см. с. 269).

Таким образом, 60% общей вариации дневной выработки объясняется вариацией уровня механизации труда и тарифных разрядов бетонщиков.

При линейной связи результирующего признака со всеми включенными в уравнение регрессии факторами показатели тесноты связи называются не индексами, а коэффициентами. Совокупный коэффициент детерминации $R_{02...s}^2$ имеет тот же смысл, что и совокупный индекс детерминации. Он может быть вычислен не только на основе факторной и общей дисперсий (формула 8.26), но и с помощью следующих формул:

$$R_{02...s}^2 = (a_1 r_{01} + a_2 r_{02} + \dots + a_s r_{0s}) : a_0 = \beta_1 r_{01} + \beta_2 r_{02} + \dots + \beta_s r_{0s}, \quad (8.28a, б)$$

где $a_i = a_{0(i-1)(i-1) \dots (i-1)s}$ — коэффициент уравнения множественной регрессии при факторе x_i ;
 σ_i — среднее квадратическое отклонение фактора x_i ;

r_{0i} — парный коэффициент корреляции;

$\beta_i = a_i \sigma_i : \sigma_0$ — бета-коэффициент регрессии, выраженный в стандартизованном масштабе). (8.29)

Определим R^2 для нашего примера, где $a_1 = 0,0356$; $a_2 = 0,8351$ (см. с. 278); $\sigma_0 = 1,3263$; $\sigma_1 = 17,1674$; $\sigma_2 = 0,7211$; $r_{01} = 0,6605$;

$r_{02} = 0,6566$ (см. с. 271). Бета-коэффициенты равны: $\beta_1 = a_1\sigma_1 : \sigma_0 = 0,0356 \times 17,1674 : 1,3263 = 0,4608$; $\beta_2 = 0,8351 \times 0,7211 : 1,3263 = 0,4540$. Отсюда $R_{02}^2 = 0,4608 \times 0,6605 + 0,4540 \times 0,6566 = 0,602$ (расхождение с R_{02}^2 вызвано округлениями при расчетах).

В частном случае (для двухфакторной регрессии) совокупный коэффициент может быть вычислен через парные:

$$R_{02}^2 = (r_{01}^2 + r_{02}^2 - 2r_{01}r_{02}r_{12}) : (1 - r_{12}^2). \quad (8.10)$$

Так, в нашем примере получим: $R_{02}^2 = (0,4363 + 0,4311 - 2 \times 0,6605 \times 0,6566 \times 0,4394) : (1 - 0,1931) = 0,603$.

Совокупный коэффициент корреляции (R), в отличие от совокупного индекса корреляции (I), может принимать как положительные значения (прямая линейная связь), так и отрицательные (обратная линейная связь), тогда как равенство его нулю говорит об отсутствии линейной связи.

Бета-коэффициенты, используемые в формуле (8.26), имеют при анализе модели и самостоятельное значение. Дело в том, что коэффициенты множественной регрессии [$a_{01(23...n)}$ и т. д.] зависят от единиц измерения соответствующих факторов и поэтому непосредственно несопоставимы между собой. Более того, даже если все факторы измеряются в одинаковых единицах, коэффициенты регрессии при них все же несравнимы, так как не учитывают степени вариации факторов. Чтобы сделать коэффициенты регрессии сопоставимыми, все признаки нужно выразить в среднеквадратических отклонениях (в стандартизованном масштабе): $t_i = \frac{x_i - \bar{x}_i}{\sigma_i}$.

Параметры уравнения регрессии в этом случае тоже будут выражены в стандартизованном масштабе и будут представлять собой бета-коэффициенты, связанные с обычными параметрами указанными выше соотношением (8.29). Бета-коэффициенты (β_i) показывают, на сколько средних квадратических отклонений σ_0 изменяется x_0 при увеличении x_i на одно среднее квадратическое отклонение σ_i и при неизменности остальных факторов, входящих в уравнение регрессии. Для парной линейной регрессии $\beta = r$.

В рассматриваемом примере уравнения множественной регрессии параметры $a_{01(2)}$ и $a_{02(1)}$ значительно отличаются по величине от другого, тогда как β -коэффициенты почти равны. Это свидетельствует о том, что степень влияния обоих факторов на результирующий признак примерно одинакова: изменение любого из факторов на одно среднеквадратическое отклонение сопровождается изменением x_0 примерно на 0,46 σ_0 .

Для множественной регрессии могут также быть найдены частные коэффициенты эластичности x_0 относительно каждого отдельного фактора:

$$\mathcal{E}_1 = \frac{\partial x_{012...n}}{\partial x_1} \cdot \frac{x_1}{x_{012...n}}, \quad (8.31)$$

$\frac{\partial x_{012...n}}{\partial x_i}$ — первая частная производная $\hat{x}_{012...n}$ по x_i .

Частные коэффициенты эластичности показывают, на сколько процентов изменяется в среднем x_0 при увеличении x_i на 1% и при фиксировании других факторов на том или ином уровне.

В нашем примере $\mathcal{E}_1 = 0,0356x_1 : (0,5002 + 0,0356x_1 + 0,8351x_2)$. При фиксировании обоих факторов на среднем уровне ($x_1 = 56$; $x_2 = 3$) получим средний коэффициент эластичности: $\mathcal{E}_1 = 0,40$. Следовательно, при увеличении x_1 с 56% до 56,56% (в 1,01 раза) x_0 увеличивается на 0,4%, если x_2 зафиксирован на среднем уровне. Аналогичные расчеты для x_2 дают $\mathcal{E}_2 = 0,50$.

Если в уравнении множественной регрессии все факторы, кроме одного, зафиксировать на каком-либо (например, на среднем) уровне, то получим уравнение условно-чистой регрессии. Так, в нашем примере, подставив в уравнение множественной регрессии $x_2 = x_2 = 3$, получим: $\hat{x}_{01(2)} = 0,5002 + 0,0356x_1 + 0,8351 \times 3 = 3,0055 + 0,0356x_1$. Уравнения условно-чистой регрессии могут быть изображены графически.

В случае множественной регрессии рассмотренные выше показатели тесноты связи (парные и совокупный индексы детерминации и корреляции) могут быть дополнены индексами частной детерминации и корреляции. Подобно тому как коэффициент вариации множественной регрессии характеризует условно-чистое значение соответствующего фактора на x_0 при исключении влияния других факторов, входящих в уравнение регрессии, так и частный индекс детерминации характеризует тесноту связи между x_0 и данным фактором при элиминировании связи x_0 с другими факторами, включенными в уравнение регрессии.

Для парной регрессии, характеризующей зависимость дневной выработки (x_0) от уровня механизации труда (x_1), выше была вычислена факторная дисперсия $\sigma_0^2 = 0,7660$ (см. с. 269). После того как в уравнение регрессии был включен второй фактор — тарифный разряд (x_2), факторная дисперсия увеличилась до $\sigma_{02}^2 = 1,0554$, т. е. на 0,2894 ($\sigma_{02}^2 - \sigma_0^2$). Следовательно, это увеличение вариации x_0 , объясняемое входящими в уравнение регрессии факторами, и представляет собой ту часть вариации, которая объясняется новым фактором x_2 , дополнительно включенным в модель. Иными словами, прирост факторной дисперсии ($\sigma_{02}^2 - \sigma_0^2$) характеризует вариацию x_0 , которую удалось дополнительно объяснить, включив в модель фактор x_2 . Если этот прирост отнести к остаточной дисперсии для парной регрессии ($\sigma_{01}^2 = \sigma_0^2 - \sigma_{01}^2$), т. е. к вариации, объясняемой прочими, не входящими в парную регрессию факторами, начиная с x_2 , то мы получим частный индекс детерминации между x_0 и x_2 при элиминировании влияния x_1 :

$$I_{02(1)}^2 = \frac{\sigma_{02}^2 - \sigma_{01}^2}{\sigma_{01(1)}^2} = \frac{\sigma_{02}^2 - \sigma_{01}^2}{\sigma_0^2 - \sigma_{01}^2}. \quad (8.32a, б)$$

Этот индекс характеризует тесноту связи между x_0 и x_2 при исключении влияния на эти признаки фактора x_1 , т. е. при условии, что вариация x_0 и x_2 под влиянием x_1 устранена, элиминирована. Если вариацию x_0 , которая объясняется факторами, включенными в уравнение регрессии, называть объясненной вариацией, а остаточную вариацию, которая этими факторами не объясняется, — не объясненной, то смысл частного индекса детерминации можно выразить так: этот индекс показывает, какая часть необъясненной ранее вариации можно объяснить, включив в модель данный фактор.

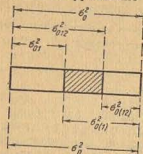


Рис. 8.7. Общая (σ_0^2), факторные (σ_{01}^2 и σ_{012}^2) и остаточные ($\sigma_{0(1)}^2$ и $\sigma_{0(12)}^2$) дисперсии для парной и множественной регрессии

т. е. 29,1% вариации x_0 необъясненной парной регрессией объясняется фактором x_2 .

Для фактора x_1 будем иметь:

$$i_{01(2)}^2 = \frac{\sigma_{012}^2 - \sigma_{02}^2}{\sigma_0^2 - \sigma_{02}^2} = \frac{1,0554 - 0,7589}{1,7592 - 0,7589} = 0,296.$$

Выше (см. с. 270) был вычислен парный индекс детерминации $i_{01}^2 = 0,435$. Поскольку фактор x_1 связан с x_2 , парный индекс, кроме влияния x_1 , отражает и косвенное влияние x_2 . Поэтому элиминирование влияния x_2 привело к уменьшению частного индекса детерминации по сравнению с парным.

В общем виде при элиминировании влияния факторов, начиная с x_1 и кончая x_{s-1} , формула частного индекса детерминации между x_0 и x_s имеет такой вид:

$$i_{0s(12 \dots s-1)}^2 = \frac{\sigma_{012 \dots s-1}^2 - \sigma_{012 \dots s-1}^2}{\sigma_0^2 - \sigma_{012 \dots s-1}^2} = \frac{\sigma_{02(1) \dots s-1}^2 - \sigma_{02(1) \dots s-1}^2}{\sigma_0^2 - \sigma_{012 \dots s-1}^2}. \quad (8.33а, б)$$

Сравнивая формулу (8.33б) с формулой совокупного индекса детерминации при s факторах (8.26), нетрудно видеть, что элиминирование влияния факторов достигается путем вычитания из числителя и знаменателя совокупного индекса детерминации факторной дисперсии, объясняемой факторами, влияние которых нужно элиминировать.

Квадратный корень из частного индекса детерминации называется частным индексом корреляции $i_{0s(12 \dots s-1)}$. Частные ин-

дексы детерминации и корреляции изменяются от 0 до 1. При линейной форме связи они называются частными коэффициентами детерминации и корреляции. Частный коэффициент детерминации $r_{0s(12 \dots s-1)}^2$ имеет такой же смысл, как и частный индекс детерминации, но применительно к линейной связи. Частные коэффициенты корреляции $r_{0s(12 \dots s-1)}$ имеют те же знаки, что и соответствующие им коэффициенты множественной линейной регрессии $a_{0s(12 \dots s-1)}$.

При двухфакторной линейной регрессии, как это было в нашем примере, частные коэффициенты корреляции могут быть вычислены через парные коэффициенты корреляции:

$$r_{01(2)} = (r_{01} - r_{02} \cdot r_{12}) : \sqrt{(1 - r_{02}^2)(1 - r_{12}^2)}, \quad (8.34а, б)$$

$$r_{02(1)} = (r_{02} - r_{01} \cdot r_{12}) : \sqrt{(1 - r_{01}^2)(1 - r_{12}^2)}.$$

Показатели множественной регрессии и корреляции, как и показатели парной, могут оказаться подверженными действию случайных факторов. Проверка их значимости (существенности), т. е. оценки того, насколько они свободны от случайных воздействий, производится с помощью критерия t Стьюдента и критерия Фишера.

При линейной зависимости x_0 от двух факторов критерий t , используемый для оценки значимости коэффициентов регрессии, рассчитывается так:

$$t_1 = a_{01(2)} \cdot \sqrt{(n-3)(1 - r_{12}^2)} : \sigma_{0(12)},$$

$$t_2 = a_{02(1)} \cdot \sqrt{(n-3)(1 - r_{12}^2)} : \sigma_{0(12)},$$

где $r_{12}^2 = r_{21}^2$ — коэффициент детерминации, характеризующий тесноту связи между x_1 и x_2 .

Вычисленные по этим формулам значения t сравниваются с критическими (табличными) значениями при принятом уровне значимости α и числе степеней свободы $k = n - m$.

Для построенного выше уравнения множественной регрессии имеем: $n = 50$; $m = 3$; $a_{01(2)} = 0,0356$; $a_{02(1)} = 0,8351$; $\sigma_1 = 17,1674$; $\sigma_2 = 0,7211$; $r_{12}^2 = r_{21}^2 = 0,1931$; $r_{12} = 0,8389$. Отсюда $t_1 = 0,0356 \times 17,1674 \cdot \sqrt{(50-3)(1-0,1931)} : 0,8389 = 4,49$; $t_2 = 4,42$. По таблице распределения Стьюдента для $k = 47$ и $\alpha = 0,05$ находим $t_{кр} = 2,01$. Так как вычисленные значения t_1 и t_2 больше критического, оба коэффициента $|a_{01(2)}|$ и $|a_{02(1)}|$ значимы.

Значимость совокупных индексов и коэффициентов детерминации (и корреляции) оценивается с помощью критерия F , который рассчитывается по формуле, приведенной на с. 273, в которой факторная и остаточная дисперсии берутся для множественной регрессии. Так, в нашем примере $\sigma_{012}^2 = 1,0554$, $\sigma_{0(12)}^2 = 0,7038$,

откуда $F = \frac{1,0554}{0,7038} \times \frac{50-3}{3-1} = 35,24$. $F_{кр}$ при $\alpha = 0,05$ ($k_1 = 2$; $k_2 = 47$) равно 3,89. Связь x_0 с совокупностью двух факторов существенна.

Оценка значимости коэффициентов регрессии с помощью критерия t часто используется для завершения отбора факторов в процессе шагового анализа. Он заключается в том, что после решения модели и оценки значимости всех коэффициентов регрес-

сии из модели исключается тот фактор, коэффициент при котором незначим и имеет наименьший коэффициент доверия t . После этого модель решается заново и снова производится оценка значимости всех оставшихся коэффициентов регрессии. Если среди них опять окажутся незначимыми, то снова исключается фактор с наименьшим коэффициентом t . Процесс исключения факторов продолжается до тех пор, пока не будет получено уравнение регрессии, все коэффициенты в котором значимы.

Шаговой анализ целесообразен в тех случаях, когда основной интерес представляют сами коэффициенты регрессии как показатели эффекта влияния факторов. Если же задачи исследования предусматривают главным образом использование модели и целью для расчета теоретических значений \hat{y} , то недостаточный уровень значимости коэффициента регрессии не является еще решающим аргументом в пользу исключения из модели соответствующего фактора, особенно если он важен экономически и его нельзя игнорировать в практическом воздействии на результативный показатель. Поэтому нередко из модели исключаются лишь те факторы, без которых существенно не увеличивается скорректированная остаточная дисперсия (при ее расчете сумма квадратов отклонений делится не на n , а на $(n-m)$, т. е. на число наблюдений за вычетом числа параметров уравнения регрессии).

Применение регрессионно-корреляционных моделей

Регрессионно-корреляционные модели могут быть использованы для решения различных задач: для анализа уровней социально-экономических явлений и процессов, например, для анализа хозяйственной деятельности предприятий и вскрытия резервов, для прогнозирования и различных плановых расчетов.

Использование моделей позволяет значительно расширить возможности анализа, в частности анализа хозяйственной деятельности предприятий. При обычных, традиционных методах анализа и оценки деятельности предприятий показатели данного предприятия сравниваются со среднеотраслевыми (в сельском хозяйстве — со среднерайонными) показателями. Такие сравнения основаны на допущении, что все предприятия отрасли, работа на той или иной группе работают примерно в одинаковых условиях, располагают более или менее одинаковыми объективными возможностями. Однако далеко не всегда дело обстоит именно так. Уровень результативного показателя на данном предприятии и средний его уровень по той или иной совокупности предприятий отличаются один от другого за счет очень многих факторов. Одни из них поддаются регулированию (управлению), целиком и полностью зависят от предприятия и отражают качество его работы, другие же лишь частично зависят от предприятия или даже почти не зависят от него (например, качество пашни в колхозе и т. п.).

Более обоснованно и правильно поэтому сравнивать фактический уровень результативного показателя (признака) на данном

предприятии не со средним, а с расчетным уровнем, который вычислен по модели путем подстановки в нее значений факторов на данном предприятии. В этом случае сопоставляемые уровни (\hat{y}_0 и \bar{y}_0) будут отличаться один от другого только за счет факторов, не входящих в модель. Следующим же этапом анализа будет сравнение расчетного (\hat{y}_0) и среднего (\bar{y}_0) уровней, разность которых обусловлено только факторами, включенными в модель. В результате отклонение фактического уровня от среднего ($x_0 - \bar{x}_0$) распадается на две части: ($x_0 - \hat{x}_0$) и ($\hat{x}_0 - \bar{x}_0$).

Если в модель входят наиболее важные и существенные факторы, охватывающие основной комплекс объективных условий, от которых зависит уровень результативного показателя (признака), то остальные, не включенные в модель факторы, характеризуют главным образом индивидуальные, субъективные условия (например, организационные факторы, не поддающиеся количественной характеристике). В таких случаях общее влияние факторов, не включенных в модель, условно, с известными оговорками можно рассматривать как результат более или менее эффективного использования объективных условий и возможностей, т. е. как эффективность использования факторов, входящих в модель (оговорка заключается в том, что подразделить факторы на объективные и субъективные можно лишь с известной условностью). Поэтому расчетные уровни для отдельных предприятий (или для каких-либо их групп) выражают такие уровни результативного признака, которые были бы достигнуты при данных (фактических) значениях факторов, входящих в модель и при среднем по всей совокупности эффективности их использования.

Сравнение же фактического уровня с расчетным ($x_0 - \hat{x}_0$ либо $\hat{x}_0 - \bar{x}_0$) показывает, насколько эффективнее использование факторов на данном предприятии выше или ниже по сравнению со средней эффективностью их использования (если результативными признаками являются производительность труда, урожайность и т. д., то более высокая эффективность имеет место при ($x_0 - \hat{x}_0 > 0$), если же результативными признаками являются себестоимость или трудоемкость единицы продукции, то при ($x_0 - \hat{x}_0 < 0$)).

Предприятия, у которых эффективность использования включенных в модель факторов значительно выше средней, могут быть исследованы с целью выявления передового опыта (или каких-либо благоприятных обстоятельств, способствовавших высокой эффективности), предприятия же, у которых эффективность существенно ниже средней, необходимо исследовать, чтобы выявить причины этого и вскрыть неиспользованные резервы.

Предприятия, у которых эффективность использования включенных в модель факторов ниже средней, имеют резервы, связанные с возможным повышением уровня эффективности. Размер этих резервов зависит от того, ставится ли задача подтянуть эти

предприятия до среднего по всей совокупности уровня эффективности использования факторов или же до того уровня, который уже достигнут остальными предприятиями. Резервы, связанные с достижением среднего уровня эффективности, можно определить, подставив в модель средние значения факторов на предприятии, где эффективность ниже средней, и найдя затем отклонение среднего фактического уровня результативного признака на этих предприятиях от расчетного уровня: $\bar{x}_0 - \hat{x}_0$.

Определим эти резервы на основе модели, выражающей зависимость дневной выработки от уровня механизации труда и тарифного разряда бетонщиков: $\hat{x}_{012} = 0,5002 + 0,0356x_1 + 0,8351x_2$. Рассчитанные по этой модели уровни выработки ниже фактических у 24 рабочих (1,5—9,13 и т. д.). У этих рабочих $x_1 = 58,4$, $x_2 = 3$, откуда $\hat{x}_{012} = 0,5002 + 0,0356 \times 58,4 + 0,8351 \times 3 = 5,08$. Фактически же у этой группы рабочих средняя выработка составляет $\bar{x}_0 = 4,38$ м³. Следовательно, если бы эффективность использования факторов x_1 и x_2 у рабочих данной группы достигла среднего уровня, то их выработка повысилась бы на 0,7 м³ (5,08—4,38), или на 16% (0,7:4,38=1,16). Общая выработка увеличилась бы за счет этого на 16,8 м³ (0,7×24) и при прочих равных условиях составила бы 266,8 м³ или 5,34 м³ на одного рабочего, что на 6,8% выше, чем фактическая средняя выработка (5 м³).

Более полно можно учесть резервы, если иметь в виду подтягивание отстающих предприятий, рабочих и т. п. до того уровня эффективности использования факторов, который достигнут не отстающими, передовыми в этом отношении предприятиями или рабочими. В этом случае нужно выделить группу таких единиц, у которых фактический уровень результативного признака либо вообще выше (лучше) расчетного, либо выше его не менее, чем на установленное число процентов. На основе первичных данных по такой группе единиц (предприятий, рабочих) нужно затем построить новую «прогрессивную» модель. Подставив в эту новую модель средние значения факторов по остальным, т. е. отстающим и непередовым предприятиям (рабочим), можно найти новые расчетные уровни, сравнение которых с фактическими и покажет величину резервов (следует, однако, отметить, что некоторые теоретические вопросы построения «прогрессивных» моделей изучены еще недостаточно).

В нашем примере по 26 рабочим, у которых эффективность использования факторов выше средней, имеем: $\Sigma x_1 = 1398$; $\Sigma x_2 = 78$; $\Sigma x_0 = 144,9$; $\Sigma x_0^2 = 80,450$; $\Sigma x_1 x_2 = 4302$; $\Sigma x_1^2 = 244$; $\Sigma x_0 x_1 = 8096,9$; $\Sigma x_0 x_2 = 446,9$; $\Sigma x_0^3 = 834,53$. Используя эти данные и решая систему нормальных уравнений (8,24), найдем: $a_0 = 1,0094$; $a_1 = 0,0423$; $a_2 = 0,7633$, откуда «прогрессивная» модель такова: пр. $\hat{x}_{012} = 1,0094 + 0,0423x_1 + 0,7633x_2$ ($R^2 = 0,824$). Прогрессивный расчетный уровень для остальных 24 рабочих равен: пр. $\hat{x}_{012} = 1,0094 + 0,0423 \times 58,4 + 0,7633 \times 3 = 5,77$, против 4,38 фактически.

Следовательно, если бы у этих 24 рабочих эффективность использования факторов повысилась до уровня, достигнутого 26 передовыми рабочими, то выработка возросла бы на 1,39 м³ (5,77—4,38), или на 31,7%.

Подставив в новую модель средние значения факторов у всех 50 рабочих ($\bar{x}_1 = 56$; $\bar{x}_2 = 3$), найдем, что выработка составила бы в среднем 5,67 м³, что на 13,4% выше фактической средней выработки (5 м³).

Анализ отклонений расчетных уровней от общего среднего уровня (\bar{x}_0) дает возможность определить, за счет каких факторов и насколько расчетный уровень выше или ниже среднего уровня. Такой анализ может быть сделан как для отдельных единиц изучаемой совокупности, так и для тех или иных их групп. Возьмем, например, первого по списку рабочего (см. табл. 8,12).

Уровень механизации его труда ($x_1 = 35\%$) ниже среднего ($\bar{x}_1 = 56\%$) на 21%. При средней эффективности использования факторов это должно было привести к снижению выработки по сравнению со средней на $0,748$ м³ [$(x_1 - \bar{x}_1) \times a_{01(2)} = (35 - 56) \times 0,0356 = -0,748$]. За счет более низкого, чем средней, тарифного разряда выработка должна была снизиться еще на $0,835$ м³ [$(x_2 - \bar{x}_2) \times a_{02(1)} = (2 - 3) \times 0,835$]. Следовательно, при фактических значениях факторов и средней эффективности их использования выработка была бы на $1,583$ м³ ниже средней (5 м³), т. е. составила бы $3,417$ м³. Это и есть расчетная выработка \hat{x}_{012} . Фактически же выработка первого рабочего составила только $3,0$ м³, что свидетельствует о более низкой эффективности использования факторов ($x_0 : \hat{x}_0 = 3,0 : 3,417 = 0,88$, или 88%).

Об экспериментальном использовании регрессионно-корреляционных моделей для анализа и оценки деятельности колхозов и совхозов, в частности для подведения итогов соревнования, рассказывается в интересной статье В. Кокашиского «Формула без неизвестных», опубликованной в «Литературной газете» № 20 от 16 мая 1974 г.

Регрессионно-корреляционные модели находят также применение для определения плановых нормативов. Так, для предприятий одного из главков были построены модели, выражающие зависимость численности инженерно-технических работников и служащих от наиболее важных факторов. Например, численность ИТР и служащих, занятых планированием и организацией труда и заработной платы, была выражена таким уравнением: $x_{012} = 0,019x_1^{0,72}x_2^{0,11}$, где x_1 — численность промышленно-производственного персонала, x_2 — количество норм выработки, действующих в основном производстве. Расчетные значения, полученные на основе таких моделей, и использовались в качестве укрупненных плановых нормативов.

Широкое применение находят уравнения регрессии и для других плановых расчетов, в том числе для прогнозирования.

1. Ряды динамики и их виды

Понятие о рядах динамики Явления общественной жизни, изучаемые статистикой, находятся в непрерывном изменении и развитии. С течением времени от месяца к месяцу, от года к году — изменяются численность населения и его состав, объем производимой продукции, уровень производительности труда и т. д. Поэтому одной из важнейших задач статистики является изучение изменения общественных явлений во времени, т. е. изучение процесса их развития, их динамики. Эту задачу статистика решает путем построения и анализа рядов динамики.

Рядом динамики называется ряд числовых значений статистического показателя, характеризующих изменение общественного явления во времени.

Примерами динамических рядов могут служить следующие данные

Численность населения и число родившихся в СССР

Таблица 91

Годы	Численность населения (оценка на начало года; млн. человек)	Число родившихся (тыс.)
1950	178,5	4 805
1960	212,4	5 341
1965	229,6	4 253
1970	241,7*	4 226
1971	243,9	4 372
1972	246,3	4 404
1973	248,6	4 386

* По переписи на 15 января.

Каждое отдельное числовое значение показателя, характеризующее величину явления, его размер, называется уровнем.

Кроме уровней каждый ряд динамики содержит указания о тех периодах либо моментах времени, к которым относится уровень ряда. В приведенном примере каждый уровень одного ря-

да показывает, какова была численность населения СССР по состоянию на 1 января соответствующего года, а каждый уровень уровня ряда — сколько детей родилось в течение данного года.

Ряды динамики дают материал для анализа развития социально-экономических явлений. Поэтому в любой экономической работе можно найти многочисленные примеры использования и анализа динамических рядов.

Виды рядов динамики В зависимости от сущности явления и характера отображающего это явление показателя каждый уровень ряда динамики может относиться либо к тому или иному моменту времени, либо к какому-то промежутку (интервалу) времени. В соответствии с этим ряды динамики делятся на моментные и интервальные.

В результате подсчета итогов статистического наблюдения получают абсолютные показатели двух видов. Одни из них характеризуют состояние явления на тот или иной момент времени; другие — на этот момент каких-либо единиц совокупности или наличие определенного объема признака. К таким показателям относятся численность населения, парк автомобилей (т. е. их наличие), поголовье скота, жилищный фонд, товарные запасы и т. п. Величину таких показателей можно определить непосредственно только по состоянию на тот или иной момент времени. Поэтому такие показатели и соответствующие ряды динамики называются моментными.

Другие показатели характеризуют итоги какого-либо процесса за тот или иной период (интервал) времени (сутки, месяц, квартал, год, пятилетку и т. п.). Такими показателями являются, например, число родившихся, количество произведенной продукции, ввод в действие жилых домов, фонд заработной платы и т. п. Эти показатели имеют уже иной характер, и их величину можно определить только за какой-либо период (промежуток, интервал) времени. Поэтому такие показатели и соответствующие ряды динамики называются интервальными.

В табл. 91 ряд динамики численности населения является моментным, так как число жителей, исчисленное по состоянию на начало соответствующего года, т. е. на определенный момент времени, представляет собой моментный показатель. Динамический ряд числа родившихся является интервальным, так как число родившихся детей подсчитано за соответствующий год и представляет собой интервальный показатель.

Из различного характера интервальных и моментных абсолютных показателей вытекают некоторые особенности (свойства) уровней соответствующих рядов динамики.

В интервальном ряду величина уровня, представляющего собой итог какого-либо процесса за тот или иной период (интервал) времени, зависит от продолжительности этого периода (длины интервала). При прочих равных условиях уровень интервального ряда тем больше, чем больше длина интервала, к которому

относится этот уровень. Так, число родившихся за год больше чем за какой-либо квартал или месяц этого года.

Иначе обстоит дело в моментных рядах динамики. Здесь тоже есть интервалы — промежутки времени между теми датами, к которым относятся уровни ряда. Однако величина того или иного конкретного уровня здесь не зависит от длины интервала. Так, численность населения на начало 1972 г. не зависит, конечно, от того, приурочены ли другие уровни к началу каждого года, как это сделано в табл. 9.1, или же к началу каждого квартала.

Каждый уровень интервального ряда представляет собой сумму уровней за более короткие промежутки времени. Поэтому в интервальном ряду динамики уровни за последовательные (примыкающие друг к другу) периоды времени можно суммировать, получая итоги (уровни) за более длительные периоды.

Иногда путем последовательного сложения уровней интервального ряда строится ряд нарастающих итогов. В таком ряду каждый уровень представляет собой итог не только за данный отчетный период, но и за все предыдущие периоды, начиная с определенной даты (с начала года, с начала пятилетки и т. п.). Нарастающие итоги нередко приводятся в отчетах предприятий.

Построение ряда нарастающих итогов приведено в следующей таблице

Таблица 9.2

ЖИЛИЩНОЕ СТРОИТЕЛЬСТВО В СССР

(млн. кв. м общей (полезной) площади)

Годы	Построено (введено в действие)	
	в данном году	с начала девятой пятилетки
1971	107,6	107,6
1972	106,7	107,6 + 106,7 = 214,3
1973	110,3	214,3 + 110,3 = 324,6

В моментном динамическом ряду один и те же единицы совокупности обычно входят в состав нескольких уровней. Так, большая часть жителей СССР, учтенных переписью на 15 января 1970 г., входила в состав населения и на 1 января 1971 г. и на 1 января 1972 г. Поэтому при суммировании уровней моментного ряда в большинстве случаев (в частности, в нашем примере) одни единицы совокупности войдут в итог дважды, другие — трижды и большее число раз. В связи с этим суммирование уровней моментного ряда динамики само по себе не имеет смысла, так как получающиеся при этом итоги лишены самостоятельной экономической значимости.

Все рассмотренные выше динамические ряды были рядами динамики абсолютных величин. Такие ряды являются исходными и первоначальными, так как лежащие в их основе аб-

солютные показатели получаются непосредственно при подсчете данных статистического наблюдения.

Кроме рядов динамики абсолютных величин, могут быть построены также динамические ряды относительных и средних величин. Эти ряды также могут быть либо моментными, либо интервальными. Так, если на конец каждого года исчислить удельный вес поголовья коров в колхозах в общем поголовье коров в стране, получим моментный ряд динамики относительных величин. Если же за ряд лет исчислить удельный вес колхозов в общем производстве молока, то получим интервальный ряд динамики относительных величин.

В интервальных рядах динамики относительных и средних величин непосредственное суммирование уровней само по себе лишено смысла, так как относительные и средние величины являются производными и исчисляются путем деления других величин.

При построении и анализе ряда динамики нужно прежде всего обратить внимание на то, чтобы уровни ряда были сопоставимы между собой, так как только в этом случае динамический ряд будет правильно отражать процесс развития явления. Сопоставимость уровней динамического ряда — это важнейшее условие обоснованности и правильности полученных в результате анализа выводов (см. п. 3 гл. V).

При построении динамического ряда вопрос о сопоставимости его уровней имеет особое значение, потому что ряд должен охватывать большой период времени, в течение которого могли произойти изменения, нарушающие сопоставимость (территориальные изменения, изменение круга охвата объектов, методологии расчетов и т. д.).

ПОКАЗАТЕЛИ АНАЛИЗА ДИНАМИКИ

Задача статистики заключается в том, чтобы путем анализа динамики выделить однородные этапы развития того или иного явления, вскрыть и охарактеризовать существенные им закономерности, тенденции и их специфические особенности.

В процессе анализа динамики рассчитывают и используют аналитические и обобщающие производные ряды. К ним относятся абсолютный прирост, темпы роста, среднегодовой прирост, ускорения и др.

Расчет большей части этих показателей производится между номерами уровней (t = 1970—1965 = 5).

Абсолютный прирост выражает абсолютную скорость роста. Однако более полную характеристику процесса роста можно получить тогда, когда абсолютные величины дополняются относительными. Относительными показателями динамики являются темпы роста и темпы прироста, характеризую-

Рис. 9.1. Схема сравнения

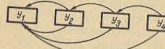


Рис. 9.1. Схема сравнения уровней при расчетах цепных и базисных показателей динамики

Абсолютный прирост Наиболее простым показателем анализа является абсолютный прирост уровня. Абсолютный прирост показывает, на сколько единиц увеличился (или уменьшился) уровень по сравнению с базисным, т. е. за тот или иной промежуток времени.

Абсолютный прирост равен разности между сравниваемым и базисным уровнями и выражается в тех же единицах, в которых измерены эти уровни:

$$\Pi = y_t - y_{t-t} \quad (9.1)$$

где P — абсолютный прирост за t единиц времени;
 y_1 — сравниваемый уровень.

1970 г., входивших в состав базисный уровень.

января 1972 г. Поэтому в каждом случае принимается пред-
явля в большинстве случаев получающихся при этом цепных аб-
единицы совокупности войдут как:
и большее число раз. В связи с этим
моментного ряда динамики само по себе не
как получающихся при этом итоге лишены самобытности, то абсо-
лютистической значимости.

Все рассмотренные выше динамические ряды аб-
динамики абсолютных величин. Такие ряды являет абсолют-
ходными, первоначальными, так как лежащие в их

РАСЧЕТ АБСОЛЮТНЫХ ПРИРОСТОВ ПРОИЗВОДСТВА
ЭЛЕКТРОЭНЕРГИИ В СССР

Таблица 9.3

Год	Произведено электроэнергии (млрд. кВт·ч)	Абсолютный прирост (млрд. кВт·ч)	
		по сравнению с предыдущим годом (цепные приросты)	по сравнению с 1970 г. (базисные приросты)
1970	741	...	×
1971	800	800 - 741 = 59	800 - 741 = 59
1972	857	857 - 800 = 57	857 - 741 = 116
1973	915	915 - 857 = 58	915 - 741 = 174

Цепные и базисные абсолютные приросты связаны между собой: сумма последовательных цепных приростов равна соответствующему базисному приросту, т. е. общему приросту за весь соответствующий промежуток времени. Так, $P_{1971} + P_{1972} + P_{1973} = 59 + 57 + 58 = 174$ (прирост за 1971—1973 гг.).

При расчете абсолютного прироста в целом за какой-либо открытый период (например, за 1966—1970 гг., т. е. за 5 лет) для интервального показателя в формуле (9.1) в качестве y_1 принимается конечный уровень данного периода (y_{1970}), а в качестве y_2 — конечный уровень предыдущего периода (y_{1965}), так как (1970—5=1965):

$$\Pi_{1966-1970} = y_{1970} - y_{1965}.$$

Если же нужно найти абсолютный прирост за те же 5 лет (1966—1970) для моментного показателя, то в качестве y_1 следует взять уровень на конец данного периода, а в качестве y_{i-1} — уровень на конец предыдущего периода (или, что то же самое, y_1 — уровень на начало следующего периода, y_{i-1} — уровень на начало данного периода): $P_{1966-1970} = y_{\text{на конец 1970 г.}} - y_{\text{на конец 1965 г.}}$

Во всех случаях разность между номерами уровней равна длине периода, за который производят расчет ($t = 1970 - 1965 = 5$).

Темп роста Абсолютный прирост выражает абсолютную скорость роста. Однако более полную характеристику процесса роста можно получить только тогда, когда абсолютные величины дополняются величинами относительными. Относительными показателями динамики являются темпы роста и темпы прироста, характеризующие

щие относительную скорость изменения уровня, т. е. интенсивность процесса роста.

Темп роста показывает, во сколько раз увеличился сравнимый уровень по сравнению с базисным (или какую часть он составляет).

Темп роста (T_p) исчисляется путем деления сравнимого уровня (y_t) на базисный (y_{t-t}):

$$T_p = \frac{y_t}{y_{t-t}} \quad (9.1)$$

Если за базу сравнения каждый раз принимается предыдущий уровень, то получаются цепные темпы роста:

$$T_p = \frac{y_t}{y_{t-1}} \quad (9.2)$$

Как и другие относительные величины, темп роста может быть выражен не только в форме коэффициента (простого отношения), но и в процентах.

Пример расчета базисных и цепных темпов роста приведен в табл. 9.4.

Таблица 9.4

Годы	Произведено электроэнергии (мрд. кВт·ч)	Темпы роста	
		по сравнению с 1970 г. (1970 г. = 1)	по сравнению с предыдущим годом (в процентах)
1970	741	1	
1971	800	800 : 741 = 1,08	(800 : 741) × 100 = 108,0
1972	857	857 : 741 = 1,16	(857 : 800) × 100 = 107,1
1973	915	915 : 741 = 1,23	(915 : 857) × 100 = 106,8

Из данных таблицы видно, что производство электроэнергии в СССР в 1971—1973 гг. росло непрерывно и высокими устойчивыми темпами.

Как и абсолютные приросты, темпы роста сами по себе являются всегда интервальными показателями, т. е. относятся к тому или иному промежутку времени.

Между цепными и базисными темпами роста, выраженными в форме коэффициентов, существует определенная взаимосвязь (если базисные темпы роста исчислены по отношению к начальному уровню ряда динамики). Эта взаимосвязь заключается в следующем:

а) произведение последовательных цепных темпов роста равно базисному темпу роста за соответствующий период;

б) частное от деления последующего базисного темпа роста на предыдущий равно соответствующему цепному темпу роста.

Взаимосвязь легко доказать.

Если нумерацию уровней вести от базисного y_0 , т. е. принять $t = 1$ и $t - t = 0$, то

$$\frac{y_1}{y_0} \times \frac{y_2}{y_1} \times \frac{y_3}{y_2} \times \dots \times \frac{y_{t-1}}{y_{t-2}} \times \frac{y_t}{y_{t-1}} = \frac{y_t}{y_0}$$

Так, если базисным является уровень 1970 г., а конечным — 1973 г., получим:

$$\frac{y_{1971}}{y_{1970}} \times \frac{y_{1972}}{y_{1971}} \times \frac{y_{1973}}{y_{1972}} = \frac{y_{1973}}{y_{1970}} \text{ и т. п.}$$

Взаимосвязь цепных и базисных темпов роста может быть использована для перехода от одних к другим, особенно в тех случаях, когда неизвестны абсолютные уровни ряда динамики. Например, известно, что в нашей стране реальные доходы в расчете на душу населения в 1965 г. возросли по сравнению с 1960 г. на 19%, т. е. в 1,19 раза, а в 1972 г. по сравнению с 1965 г. — на 45%. Чтобы определить, как возросли реальные доходы за 12 лет, т. е. в 1972 г. по сравнению с 1960 г., нужно найти темп роста, равный $y_{1972} : y_{1960}$. Нетрудно видеть, что

$$\frac{y_{1972}}{y_{1960}} = \frac{y_{1965}}{y_{1960}} \times \frac{y_{1972}}{y_{1965}} = 1,19 \times 1,45 = 1,72.$$

Следовательно, в 1972 г. по сравнению с 1960 г. реальные доходы на душу населения увеличились на 72%.

Другой пример. Произведенный национальный доход СССР составлял (в процентах к 1960 г.): в 1970 г. — 199%, в 1972 г. — 219%. Разделив второй темп роста на первый, найдем, что в 1972 г. по сравнению с 1970 г. национальный доход увеличился в 1,1 раза:

$$\left(\frac{y_{1972}}{y_{1960}} : \frac{y_{1970}}{y_{1960}} \right) = \frac{y_{1972}}{y_{1970}} \text{ или } 219 : 199 = 1,1$$

Темп прироста и абсолютное значение 1% прироста

Темп прироста характеризует относительную величину прироста, т. е. величину абсолютного прироста по отношению к базисному уровню:

$$T_{пр} = \frac{p}{y_{t-t}} \quad (9.5)$$

где $T_{пр}$ — темп прироста за t единиц времени;

p — абсолютный прирост за то же время;

y_{t-t} — базисный уровень.

Выраженный в процентах темп прироста показывает, на сколько процентов увеличился (или уменьшился) уровень по сравнению с базисным, принятым за 100%. Как и абсолютный прирост, темп прироста может быть положительным, что свидетельствует об увеличении уровня, и отрицательным, что говорит об уменьшении уровня.

Между темпом прироста и темпом роста существует непосредственная взаимосвязь:

$$T_{\text{пр}} = \frac{P}{Y_t - Y_{t-1}} = \frac{Y_t - Y_{t-1}}{Y_{t-1}} = \frac{Y_t}{Y_{t-1}} - 1 = T_r - 1. \quad (9.6)$$

Таким образом, темп прироста всегда на единицу меньше темп роста, выраженного в форме коэффициента, или на 100% меньше темпа роста, выраженного в процентах:

$$T_{\text{пр}} (\%) = T_r (\%) - 100\%. \quad (9.7)$$

Так, в нашем примере (см. табл. 9.4) в 1972 г. производство электроэнергии возросло: по сравнению с 1971 г. — на 7,1% $[(1,071 - 1) \times 100]$, а по сравнению с 1970 г. — на 16%.

При анализе темпов развития никогда не следует упускать из виду, какие абсолютные величины — уровни и абсолютные приросты — скрываются за темпами роста и прироста. В частности, нужно иметь в виду, что при снижении (замедлении) темпов роста и прироста абсолютный прирост может возрастать. Возьмем, например, производство телевизоров в СССР.

РОСТ ПРОИЗВОДСТВА ТЕЛЕВИЗОРОВ В СССР

Таблица 9.4

Годы	Произведено телевизоров (тыс. шт.)	Темп прироста за 5 лет (%)	Абсолютный прирост за 5 лет (тыс. шт.)
1960	1726		
1965	3655	112	1929
1970	6682	83	3027

За седьмую пятилетку (1961—1965 гг.) производство телевизоров возросло на 112%, т. е. более чем в 2 раза, а за восьмую пятилетку (1966—1970) — только на 83%. Однако в абсолютном выражении прирост в восьмой пятилетке был значительно больше, чем в седьмой: более 3 млн. шт. против 1,9 млн. шт. Если в седьмой пятилетке каждый из 112% прироста в абсолютном выражении составлял 17 тыс. шт. (1929 : 112), то в восьмой пятилетке каждый из 83% прироста выражался уже величиной более 36 тыс. шт. (3027 : 83), т. е. вдвое большей, чем в предыдущей пятилетке.

Использованный здесь показатель — абсолютное значение 1% прироста — находится путем деления абсолютного прироста на выраженный в процентах темп прироста. Этот показатель можно исчислить несколько иначе. Так как за 100% всегда принимается базисный уровень, то 1% будет составлять его сотую часть:

$$A = \frac{Y_t - Y_{t-1}}{100}. \quad (9.8)$$

Так, для восьмой пятилетки получим: $A = 3655 : 100 = 36,6$ (тыс. шт.).

СРЕДНИЕ ПОКАЗАТЕЛИ ДИНАМИКИ

Общие вопросы

Исчисления средних показателей динамики

С течением времени изменяются не только уровни явлений, но и показатели динамики — абсолютные приросты и темпы развития. Поэтому для обобщающей характеристики развития, для выявления его типичных особенностей и закономерностей используются средние показатели динамики: средние уровни, средние приросты, средние темпы развития. Часто приходится прибегать к этим показателям при изучении таких явлений, уровни которых то повышаются, то понижаются. Известно, например, что в сельском хозяйстве год на год не приходится: более благоприятные годы чередуются с менее благоприятными. Поэтому при анализе роста сельскохозяйственного производства особенно целесообразно сравнивать не годовые, а средние годовые показатели, в которых яснее видна основная тенденция развития.

К расчету средних показателей динамики, в частности среднего уровня, очень часто приходится также прибегать при исчислении некоторых относительных величин для обеспечения сопоставимости величин и делителя. Например, необходимо определить, сколько молока произведено в СССР в 1974 г. в расчете на душу населения. Для этого нужно количество молока, произведенного в 1974 г., разделить на численность населения в этом же году. Но производство молока за год — это интервальный показатель, относящийся ко всему годовому промежутку времени, а численность населения — показатель моментный, относящийся к определенному моменту времени и непрерывно меняющийся на протяжении года. Иско поэтому, что численность населения на тот или иной момент времени несопоставима с показателями производства за весь год. Для обеспечения сопоставимости нужно и численность населения рассчитать в целом за год, а это можно сделать, лишь определив среднюю за год численность населения.

При исчислении средних показателей динамики необходимо иметь в виду, что к этим средним показателям и полностью относятся общие положения теории средних. Это означает прежде всего, что исчисление динамических средних имеет смысл лишь для такого периода времени, который характеризуется какими-либо общими, т. е. одинаковыми в известном отношении, условиями развития явления. Если же динамическая средняя исчислена за период, охватывающий различные этапы развития явления, т. е. за период, в течение которого условия развития явления существенно менялись, то такой средней нужно пользоваться с большой осторожностью, дополняя и корректируя ее средними за отдельные этапы развития.

Средние показатели динамики должны также удовлетворять логико-математическому требованию, которое заключается в том, что при замене средней величиной тех фактических величин, из которых исчислена средняя, не должна изменяться величина некоторого обобщающего, итогового показателя, связанного с осредняемым.

Средний уровень
ряда динамики

Метод расчета среднего уровня ряда динамики зависит прежде всего от характера показателя, лежащего в основе ряда, т. е. от вида динамического ряда.

Наиболее просто исчисляется *средний уровень интервального ряда динамики абсолютных величин*. Уровни такого ряда можно суммировать, получая итоговые уровни за более продолжительные периоды. Вполне логично поэтому исчислять средний уровень интервального ряда так, чтобы при замене фактических уровней на средней величиной не изменялся итоговый уровень за весь рассматриваемый период.

Это означает, что должно иметь место такое равенство:

$$y_1 + y_2 + \dots + y_n = \bar{y} + \bar{y} + \dots + \bar{y} = n\bar{y}.$$

Это приводит к простой средней арифметической

$$\bar{y} = \frac{\sum y}{n}, \quad (9.9)$$

где n — число фактических уровней за последовательные равные промежутки времени.

Так, сложив годовые уровни производства электроэнергии в СССР за 1971—1973 гг. (см. табл. 9.4) и разделив полученную сумму на 3, узнаем, сколько электроэнергии производилось в среднем ежегодно в этот период $(800 + 857 + 915) : 3 = 857$ (млрд кВт·ч).

Как видно из приведенного примера, средний уровень интервального ряда динамики требует, чтобы было указано: а) за какой конкретный (календарный) период исчислен *средний уровень*, б) какой интервал (промежуток) времени принят в качестве единицы времени, в расчете на которую исчислен *средний уровень*. В нашем примере за 1971—1973 гг. был исчислен среднегодовой уровень, т. е. исчисление производилось в расчете на год.

Используя те же данные, можно, однако, исчислить средний уровень за те же три года, но в расчете на другую единицу времени, например среднеквартальный или среднемесячный уровень. В этих случаях количество произведенной за 3 года электроэнергии нужно разделить на продолжительность этого периода, выраженную в кварталах или месяцах.

Таким образом, в наиболее общем виде *средний уровень интервального ряда динамики абсолютных величин* равен уровню за весь изучаемый период, деленному на продолжительность (длину) этого периода в тех или иных единицах времени:

$$\bar{y} = \frac{\sum y}{t}, \quad (9.10)$$

где t — длина периода, за который делается расчет.

Несколько сложнее обстоит дело с исчислением *среднего уровня моментного ряда динамики абсолютных величин*. Моментный по-

казатель практически может изменяться почти непрерывно, особенно если он относится к большой территории. В связи с этим очевидно, что чем более подробными и исчерпывающими данными о его изменении мы располагаем, тем более точно окажется возможным исчислить средний уровень. Более того, сам метод расчета зависит от того, насколько подробно имеющиеся у нас данные, позволяют рассмотреть различные случаи.

Первый случай. Совершенно точно средний уровень моментного показателя может быть исчислен, конечно, лишь в тех случаях, когда известны полные, исчерпывающие данные о его изменении, т. е. точно известно, каков был уровень в каждый момент, когда и на сколько он увеличивался или уменьшался. В качестве определяющего показателя и критерия правильности расчета среднего уровня в этом случае выступает общая сумма дней (месяцев, лет) наличия или функционирования всех единиц совокупности, например общее количество человеко-дней, машино-дней и т. п. Величина этого показателя должна остаться без изменения при замене фактических уровней их средней величиной.

При расчете среднего уровня интервального ряда динамики абсолютных величин определяющим показателем является, как был показано выше, уровень за весь период, равный сумме всех фактических уровней. Для моментного показателя сумма уровней сама по себе не имеет самостоятельного экономического значения и не может быть принята в качестве определяющего показателя. Только переход от учета тех или иных единиц, рассматриваемых относительно к продолжительности времени их наличия (функционирования), а следовательно учету каждого календарного дня (месяца, года) наличия или функционирования каждой отдельной единицы дает такие показатели, сумма которых имеет самостоятельное экономическое значение, определенный экономический смысл и может быть принята в качестве определяющего показателя.

Допустим, что нужно определить среднюю численность работающих членов бригады за рабочую неделю. Известно, что в понедельник работали 23 человека, во вторник — 22, а среду — 20, четверг — 21, пятницу — 24. Таким образом, за 5 рабочих дней недели было отработано 110 человеко-дней $(23 + 22 + 20 + 21 + 24)$. Следовательно, в среднем ежедневно работали 22 человека (110 человеко-дней : 5 дней).

Аналогично можно определить среднюю численность всех членов бригады, т. е. рабочих, как принимавших, так и не принимавших в отдельные дни участия в работе (так называемую среднесписочную численность рабочих). В этом случае нужно за весь интересующий нас период подсчитать общее количество человеко-дней пребывания рабочих в списке бригады (цеха, предприятия) и разделить эту сумму на число календарных дней в данном периоде.

Если при этом численность рабочих изменялась не ежедневно, как в предыдущем примере, а относительно редко, то расчет удобнее оформить в виде таблицы.

Например, по состоянию на 1 сентября в списке бригады числится 25 человек, с 11-го сентября 3 человека выбыли, а с 26-го

бригада пополнилась 6 новыми рабочими. Расчет общего количества человеко-дней приведен в табл. 9.6.

РАСЧЕТ КОЛИЧЕСТВА ЧЕЛОВЕКО-ДНЕЙ

Таблица 9.6

Календарный период, в течение которого списочное число рабочих не изменялось	Продолжительность этого периода в днях t	Списочное число рабочих в течение этого времени y	Количество человеко-дней yt
1—10 сентября	10	25	250
11—25 "	15	22	330
26—30 "	5	28	140
Итого	30	—	720

Следовательно, среднесписочное число рабочих составляет 720 человеко-дней: 30 дней = 24 человека.

Обобщая рассмотренные примеры, можно сделать вывод, что при наличии исчерпывающих данных об изменении моментного показателя средней его уровень исчисляется по формуле средней арифметической взвешенной:

$$\bar{y} = \frac{\sum yt}{\sum t}, \quad (9.11)$$

где y — уровни, сохранившиеся без изменения в течение промежутков времени t .

При этом если изменения происходят ежедневно, как это было в первом примере, то каждый из промежутков времени будет равен 1 дню.

В большинстве случаев уровень моментного показателя изменяется настолько часто, что практически оказывается нецелесообразным учитывать и отражать в рядах динамики все эти изменения. Для практических потребностей бывает достаточно учитывать уровень явления только один раз в месяц, квартал или год. Поэтому в большинстве случаев для расчета среднего уровня моментного показателя мы располагаем данными о состоянии явления только на определенных даты (моменты времени), которые могут как совпадать, так и не совпадать с моментами изменения явления. В таких случаях любой расчет среднего уровня может быть только приближенным, так как отсутствие сведений о том, как изменялся уровень в промежутках между датами, не дает возможности точно подсчитать общую сумму дней (месяцев) наличия или функционирования каждой единицы (общую сумму человеко-дней, машино-дней и т. п.). В результате расчет среднего уровня приходится делать исходя из некоторых предположений и условных допущений.

Второй случай. В наиболее простом из этих случаев нам известны данные об уровне явления по состоянию только на начало и на конец периода, за который нужно исчислить средний уровень.

Пусть, например, на 1 января 1972 г. во вновь заселяемом районе города проживало 10 тыс. человек, а к концу года численность жителей района увеличилась до 16 тыс. Нужно определить среднюю численность населения района в 1972 г.

Изобразив численность населения на начало и конец года графически, получим следующую картину (см. рис. 9.2).

Если бы численность населения в течение всего года оставалась на первоначальном уровне (10 тыс. человек), то за год жителями было бы прожито 10 тыс. человеко-лет. Графически эти человеко-лет изображались бы площадью прямоугольника $ABED$. В действительности численность населения увеличилась. Условно предположим, что это увеличение в течение года происходило равномерно (в арифметической прогрессии), изменение численности населения можно изобразить прямой BC . Тогда, очевидно, число прожитых человеко-лет изобразится площадью трапеции $ABCD$:

$$S_{ABCD} = \frac{(AB + CD)}{2} \cdot AD.$$

Чтобы найти среднюю численность населения, нужно число прожитых человеко-лет разделить на длину времени, т. е. на 1 год (графически — AD):

$$\bar{y} = \frac{S_{ABCD}}{AD} = \frac{AB + CD}{2} = \frac{y_{нач} + y_{кон}}{2}.$$

Предположение о равномерном изменении уровня в период между двумя датами приводит к расчету среднего уровня как простой средней арифметической из уровней на начало ($y_{нач}$) и на конец ($y_{кон}$) периода:

$$\bar{y} = \frac{y_{нач} + y_{кон}}{2}. \quad (9.12)$$

В нашем примере средняя численность населения в 1972 г. составляла: $y = (10 + 16) : 2 = 13$ (тыс. человек).

Третий случай имеет место тогда, когда кроме уровней на начало и на конец периода, за который нужно сделать расчет, известны также уровни на некоторые промежуточные даты, отделенные друг от друга *неравными промежутками времени*.

Допустим, что численность жителей вновь заселяемого городского района составляла (тыс. человек): на 1/1 1972 г. — 10,0, на 1/1 1972 г. — 10,8, на 1/XI 1972 г. — 14,4, на 1/1 1973 г. — 16,0.

Общая схема исчисления среднего уровня в таких случаях заключается в следующем. Сначала рассчитываются средние уровни за промежутки времени между двумя соседними датами. Этот расчет делается так, как было показано в предыдущем случае, т. е. по формуле простой средней арифметической (9.12). Затем на полученных таким путем промежуточных средних уровней (\bar{y}_i) исчисляется средняя арифметическая взвешенная, причем в каче-

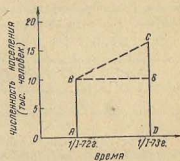


Рис. 9.2 К расчету среднего уровня по данным на начало и конец периода

стве весов принимаются величины промежутков времени между соответствующими датами:

$$\bar{y} = \frac{\sum y_i t_i}{\sum t_i}, \quad (9.13)$$

где \bar{y}_i — средние уровни за промежутки времени между соседними датами;

t_i — величины этих промежутков.

В нашем примере промежуточные средние уровни равны (тыс. человек): за январь—апрель 1972 г. — $y_{I-IV} = (10,0 + 10,8) : 2 = 10,4$; за май—октябрь — $y_{V-X} = (10,8 + 14,4) : 2 = 12,6$; за ноябрь—декабрь — $y_{XI-12} = (14,4 + 16,0) : 2 = 15,2$.

Промежутки времени между датами (t_i), измеренные в месяцах, равны: $t_1 = 4$, $t_2 = 6$, $t_3 = 2$. Следовательно, средний уровень в целом за 1972 г. составляет:

$$\bar{y} = \frac{10,4 \times 4 + 12,6 \times 6 + 15,2 \times 2}{4 + 6 + 2} = 12,3 \text{ (тыс. человек).}$$

Четвертый случай. Если кроме уровней на начало и на конец периода, за который нужно сделать расчет, известны уровни на промежуточные даты с равными интервалами времени между ними, то средний уровень можно исчислить по предыдущей формуле (9.13), однако более целесообразно преобразовать ее исходя из того, что в данном случае $t_1 = t_2 = \dots = t$:

$$\bar{y} = \frac{\sum y_i t_i}{\sum t_i} = \frac{\frac{y_1 + y_2}{2} \cdot t_1 + \frac{y_2 + y_3}{2} \cdot t_2 + \dots + \frac{y_{n-1} + y_n}{2} \cdot t_{n-1}}{t_1 + t_2 + \dots + t_{n-1}} =$$

$$= \frac{\frac{y_1 + y_2}{2} \cdot t + \frac{y_2 + y_3}{2} \cdot t + \dots + \frac{y_{n-1} + y_n}{2} \cdot t}{t(n-1)}$$

(число промежутков между датами всегда на 1 меньше числа дат и уровней).

Сократив далее дробь на t и произведя почленное деление на 2 в числителе, получим окончательно:

$$\bar{y} = \frac{\frac{1}{2} y_1 + y_2 + \dots + y_{n-1} + \frac{1}{2} y_n}{n-1}. \quad (9.14)$$

где y_1 — уровень на начало периода, за который делается расчет; y_n — уровень на конец этого периода;

n — число всех уровней на равноотстоящие даты.

В целях контроля полезно иметь в виду, что число уровней без одного, т. е. $(n-1)$, численно равно длине периода, за который делается расчет, выраженной в интервалах между датами (месяцах, кварталах и т. п.).

Использование формулы (9.14) рассмотрим на следующем примере (см. табл. 9.7).

Нужно найти средний размер остатков за 1973 г., а также за первое полугодие 1974 г.

ОСТАТКИ (ЗАПАСЫ) МАТЕРИАЛА НА СКЛАДЕ (т)

Таблица 9.7

Годы	На 1 января	На 1 апреля	На 1 июля	На 1 октября
1973	20	15	14	25
1974	24	17	16	...

Поскольку интервалы между датами равные (квартал), можно применить формулу (9.14). При расчете средних остатков за 1973 г. в качестве уровня на конец года (y_n) берем уровень на 1 января 1974 г.

$$\bar{y}_{1973} = (10 + 15 + 14 + 25 + 12) : 4 = 19 \text{ (т)},$$

$$\bar{y}_{\text{пол. 1974}} = (12 + 17 + 8) : 2 = 18,5 \text{ (т)}.$$

Формулы (9.12), (9.13) и (9.14) при прочих равных условиях дают тем более точные результаты, чем меньше колеблется уровень и чем короче промежутки времени между датами.

Средний уровень рядов динамики относительных и средних величин следует исчислять исходя из содержания и смысла этих относительных и средних показателей. Для этого относительный или средний показатель, являющийся уровнем ряда динамики, нужно представить в виде исходной базы его расчета, т. е. в виде отношения абсолютных величин. Далее по рассмотренным выше правилам и формулам нужно найти средние уровни этих абсолютных величин, а затем путем деления одного из них на другой вычислить искомый средний уровень относительного (или среднего) показателя.

Обозначив относительный (или средний) показатель, лежащий в основе ряда динамики, через $k = \frac{a}{b}$, где a и b — абсолютные величины, порядок расчета динамического среднего уровня такого показателя можно выразить формулой

$$\bar{k} = \frac{\bar{a}}{\bar{b}}. \quad (9.15)$$

Таким образом, сначала производится динамическое осреднение абсолютных величин, отношением которых является уровень ряда динамики, а затем путем деления этих средних уровней исчисляется средний уровень относительного (или среднего) показателя.

Например, нужно найти средний размер вклада за 1971—1972 гг. по следующим данным (см. табл. 9.8).

Средний размер вклада на тот или иной момент времени рассчитывается путем деления суммы вкладов на их число. Поэтому при расчете среднего размера вклада за какой-либо период нужно среднюю сумму вкладов за этот период разделить на среднее число вкладов.

ВКЛАДЫ НАСЕЛЕНИЯ В СБЕРЕГАТЕЛЬНЫЕ КАССЫ СССР
(на конец года)

Таблица 9.8

Годы	Число вкладов (млн.)	Сумма вкладов (млн. руб.)	Средний размер вклада (руб.)
1970	80 121	46599,9	581,6
1971	84 549	53215,3	629,4
1972	89 163	60732,0	681,1

Среднюю сумму и среднее число вкладов найдем по формуле (9.14):

$$\text{Средняя сумма вкладов} = \frac{0,5 \times 46599,9 + 53215,3 + 0,5 \times 60732,0}{2} = 53440,6 \text{ (млн. руб.)}.$$

$$\text{Среднее число вкладов} = \frac{0,5 \times 80121 + 84549 + 0,5 \times 89163}{2} = 84595,5 \text{ (тыс.)}.$$

Отсюда средний размер вклада за 1971—1972 гг. равен:
53440,6 : 84 595,5 = 631,7 (руб.).

Средний абсолютный прирост — на сколько единиц увеличивался или уменьшался уровень по сравнению с предыдущим в среднем за ту или иную единицу времени (в среднем ежемесячно, ежегодно и т. п.).

Средний абсолютный прирост характеризует среднюю абсолютную скорость роста (или снижения) уровня.

Абсолютный прирост всегда является интервальным показателем. Поэтому его средняя величина исчисляется по формуле простой средней арифметической из цепных приростов за последовательные и равные промежутки времени:

$$\bar{\Pi} = \frac{\sum \Pi}{t}, \quad (9.16)$$

где $\bar{\Pi}$ — средний абсолютный прирост;
 t — число цепных приростов.

Например, нужно определить, на сколько миллионов квадратных метров увеличилось в среднем ежегодно производство шелковых тканей в СССР в 1971—1973 гг. (см. табл. 9.9).

Отсюда $\bar{\Pi}_{1971-1973} = (44 + 80 + 75) : 3 = 66,3$ (млн. кв. м)

Из приведенного примера видно, что, говоря о среднем абсолютном приросте, нужно указывать: а) за какой конкретный (календарный) период нужно исчислить (или исчислен) средний прирост, б) какой интервал принят за единицу времени, в расчете на которую нужно исчислить (или исчислен) средний прирост. Так, в нашем примере за 1971—1973 гг., т. е. за 3 года, исчислен среднегодовой абсолютный прирост (средний прирост в расчете на 1 год).

Таблица 9.9

РОСТ ПРОИЗВОДСТВА ШЕЛКОВЫХ ТКАНЕЙ В СССР

Годы	Произведено (млн. кв. м)	Абсолютный прирост по сравнению с предыдущим годом
1970	1146	44
1971	1190	80
1972	1270	75
1973	1345	

В формуле (9.16) сумма цепных приростов ($\sum \Pi$) представляет собой прирост за весь календарный период в целом, а число цепных приростов (t) равно длине этого периода, выраженной в соответствующих единицах времени. Поэтому формулу среднего абсолютного прироста можно выразить через исходные данные, т. е. через уровни ряда динамики:

$$\bar{\Pi} = \frac{y_t - y_{t-1}}{t}, \quad (9.17)$$

где y_t — конечный уровень периода, за который производится расчет;

i — номер конечного уровня;

t — длина периода, за который производится расчет;

y_{t-1} — базисный уровень.

При расчете $\bar{\Pi}$ по формуле (9.17) за тот или иной период для интервального показателя в качестве y_t принимается конечный уровень периода, за который делается расчет, в качестве базисного уровня (y_{t-1}) — конечный уровень предыдущего периода. Каждый уровень в формуле (9.17) должен в этом случае относиться к той же единице времени, в расчете на которую исчисляется средний прирост.

Для контроля следует иметь в виду, что разность между номерами уровней в числителе формулы (9.17) равна длине периода, за который производится расчет: $i - (i - t) = t$.

В нашем примере (табл. 9.9) расчет производится за 3 года (1971—1973 гг.), т. е. $t = 3$ годам, $i = 1973$, $i - t = 1970$, т. е. в качестве базисного следует принять уровень 1970 г.:

$$\bar{\Pi}_{1971-1973} = \frac{y_{1973} - y_{1970}}{3} = \frac{1345 - 1146}{3} = 66,3 \text{ (млн. кв. м)}.$$

Если показатель (уровень ряда динамики) является моментным, то в качестве y_t принимается уровень на конец периода, за который делается расчет, а в качестве базисного уровня (y_{t-1}) — уровень на конец предыдущего периода (или соответственно y_t — уровень на начало следующего периода, y_{t-1} — уровень на начало данного периода). Исчисление среднего прироста в этом случае можно производить в расчете на любую единицу времени.

Например, городской жилищный фонд СССР составлял (на конец года; млн. кв. м общей (полезной) площади жилищ): 1965 г. — 1238, 1972 г. — 1661. Определим, на сколько миллионов кв. м увеличился жилищный фонд в 1966—1972 гг. в среднем ежегодно.

$$\overline{T}_{1966-1972} = \frac{y_{\text{кон. 1972}} - y_{\text{кон. 1965}}}{7} = \frac{1661 - 1238}{7} = 60,4 \text{ (млн. кв. м)}.$$

В данном случае длина периода, за который сделан расчет, выражена в годах и равна разности между номерами уровней в числителе формулы. Если длину периода измерять в кварталах, то она будет равна 28 кварталам (7×4), а среднеквартальный абсолютный прирост составит:

$$\overline{T} = \frac{1661 - 1238}{28} = 15,1 \text{ (млн. кв. м)},$$

т. е. городской жилищный фонд в 1966—1972 гг. в среднем за квартал возрастал на 15,1 млн. кв. м.

Формула (9.17) более удобна для расчетов, чем формула (9.16), так как не требует вычисления всех ценных абсолютных приростов.

Средние темпы роста и прироста

В качестве основы и критерия правильности исчисления среднего темпа роста могут быть приняты различные определяющие показатели. Обычно в качестве такой основы принимается произведение ценных темпов роста, равное темпу роста за весь рассматриваемый период. Обозначив базисный (первый) уровень y_0 , т. е. приняв $i = t$ и $i - t = 0$ (см. с. 299), будем иметь:

$$\frac{y_1}{y_0} \times \frac{y_2}{y_1} \times \dots \times \frac{y_{t-1}}{y_{t-2}} \times \frac{y_t}{y_{t-1}} = T_1 \times T_2 \times \dots \times T_{t-1} \times T_t =$$

$$= \frac{y_t}{y_0} = \frac{y_i}{y_{i-t}}. \quad (\text{A})$$

При этом ставится задача найти такой средний темп роста (\overline{T}_p), чтобы при замене им фактических ценных темпов в формуле (А) остался без изменения темп роста за весь период ($y_i - y_{i-t}$). Таким образом, должны соблюдаться следующие равенства:

$$T_1 \times T_2 \times \dots \times T_{t-1} \times T_t = \frac{y_i}{y_{i-t}} = \overline{T}_p \times \overline{T}_p \times \dots \times \overline{T}_p.$$

произведение t средних темпов роста

Следовательно,

$$(\overline{T}_p)^t = T_1 \times T_2 \times \dots \times T_t = \frac{y_i}{y_{i-t}}.$$

Отсюда может быть получена формула среднего темпа роста в двух видах:

$$\overline{T}_p = \sqrt[t]{T_1 \times T_2 \times \dots \times T_t}. \quad (9.18)$$

Эта формула носит название *простой средней геометрической* и представляет собой корень степени t из произведения t ценных темпов роста.

Вторая формула является преобразованием первой:

$$\overline{T}_p = \sqrt[t]{\frac{y_i}{y_{i-t}}}, \quad (9.19)$$

где y_i — конечный уровень периода, за который производится расчет;

y_{i-t} — базисный уровень;

t — длина периода, за который делается расчет, равная разности между номерами уровней в подкоренном выражении.

Особенности применения этой формулы к интервальным и моментным показателям аналогичны специфике использования в этих случаях формулы (9.17).

Средний темп роста, выраженный в форме коэффициента, показывает, во сколько раз увеличился уровень по сравнению с предыдущим в среднем за единицу времени (в среднем ежегодно, ежемесячно и т. п.).

Для средних темпов роста и средних темпов прироста сохраняют силу та же взаимосвязь, которая имеет место между обычными темпами роста и прироста:

$$\overline{T}_{пр} = \overline{T}_p - 1,$$

или

$$\overline{T}_{пр} (\%) = \overline{T}_p (\%) - 100\%. \quad (9.20)$$

Средний темп прироста (или снижения), выраженный в процентах, показывает, на сколько процентов увеличился (или уменьшился) уровень по сравнению с предыдущим в среднем за единицу времени (в среднем ежегодно, ежемесячно и т. п.).

Средний темп прироста характеризует среднюю интенсивность роста, т. е. среднюю относительную скорость увеличения уровня.

Из двух видов формулы среднего темпа роста чаще используется формула (9.19), так как она не требует вычисления всех ценных темпов роста. По формуле же (9.18) расчет целесообразно производить лишь в тех случаях, когда не известны ни уровни ряда динамики, ни темп роста за весь период, а известны только ценные темпы роста (или прироста). Например, рост производительности общественного труда в СССР (в процентах к предыдущему году) составлял:

в 1970 г. — 107,0; в 1971 г. — 104,4; в 1972 г. — 102,8; в 1973 г. — 107,5.

Используя формулу (9.18), получим:

$$\overline{T}_p = \sqrt[4]{107,0 \times 104,4 \times 102,8 \times 107,5}.$$

Расчет производится либо с помощью логарифмов, либо по специальным таблицам¹, что более удобно.

В нашем примере $T_p = 105,4\%$, откуда $T_{np} = 5,4\%$.

Следовательно, производительность общественного труда в СССР в 1970—1973 гг. в среднем ежегодно повышалась на 5,4%.

Формула 9.19 дает возможность исчислять средний темп роста в двух случаях: 1) если известны базисный (y_{i-1}) и конечный (y_i) уровни, 2) если известен темп роста (либо темп прироста) за весь период.

Исчислим, например, предусмотренный Директивой XXIV съезда КПСС плановый среднегодовой темп роста производства минеральных удобрений в 1971—1975 гг. По пятилетнему плану в 1975 г. должно быть произведено 90 млн. т минеральных удобрений против 55,4 млн. т в 1970 г.

Используя формулу (9.19), получим:

$$t = 1975 - 1970 = 5;$$

$$\bar{T}_p = \sqrt[5]{\frac{y_{1975}}{y_{1970}}} = \sqrt[5]{\frac{90}{55,4}}; \lg \bar{T}_p = (\lg 90 - \lg 55,4) : 5 = 0,0421,$$

$$\text{откуда } \bar{T}_p = 1,102, \text{ или } 110,2\%.$$

Следовательно, производство минеральных удобрений в девятой пятилетке (1971—1975 гг.) должно увеличиваться в среднем ежегодно на 10,2%. С этой величиной можно сопоставлять годовые темпы прироста в 1971—1975 гг. Так, в первом году девятой пятилетки (1971 г.) производство удобрений возросло по сравнению с предыдущим 1970 г. на 10,8%, т. е. несколько больше, чем и среднем предусматривает план.

Второй случай использования формулы (9.19) рассмотрим на следующем примере. Известно, что в нашей стране в 1970 г. по сравнению с 1965 г. реальные доходы на душу населения возросли на 33%, т. е. в 1,33 раза. Длина периода здесь также составляет 5 лет ($t = 1970 - 1965 = 5$), а темп роста за эти 5 лет, т. е. $y_{1970} : y_{1965}$, равен 1,33.

Следовательно, среднегодовой темп роста составлял:

$$\bar{T}_p = \sqrt[5]{\frac{y_{1970}}{y_{1965}}} = \sqrt[5]{1,33} = 1,059, \text{ или } 105,9\%,$$

т. е. реальные доходы на душу населения в 1966—1970 гг. в среднем ежегодно повышались почти на 6%.

В случае если уровень ряда динамики снижается, то темп роста за весь период и средний темп роста будет меньше 1 (или 100%), а средний темп прироста — отрицательной величиной. Отрицательный \bar{T}_{np} представляет собой средний темп снижения и характеризует среднюю относительную скорость снижения уровня.

¹ См.: Айрапетов А. М. Таблицы исчисления среднегодовых темпов роста, прироста и снижения. М., «Статистика», 1971; Харламов А. И. Схемы и таблицы расчета средних темпов динамики. М., «Статистика», 1971.

Так, детская смертность в СССР снизилась с 35 на тысячу рожденных в 1960 г. до 26 в 1973 г. Следовательно, $y_{1973} : y_{1960} = 26 : 35 = 0,7429$, а длина периода $t = 1973 - 1960 = 13$. Отсюда среднегодовой темп снижения равен:

$$\bar{T}_{cn} = \sqrt[13]{\frac{y_{1973}}{y_{1960}}} - 1 = \sqrt[13]{0,7429} - 1 = 0,977 - 1 = -0,023,$$

или —2,3%, т. е. детская смертность в 1961—1973 гг. ежегодно снижалась в среднем на 2,3%.

Показатели абсолютного и относительного ускорения или замедления скорости изменения уровня

На разных этапах развития рост или снижение уровня могут происходить с разной скоростью — быстрее или медленнее. Для числовой характеристики ускорения или замедления роста (снижения) уровня можно сопоставлять одноименные показатели динамики, исчисленные за разные

этапы развития явления или за отдельные отрезки времени одного и того же этапа. Если сравниваемые этапы развития или отрезки времени имеют одинаковую продолжительность, можно сопоставлять показатели динамики за весь период. Однако чаще длина сравниваемых периодов бывает неодинакова, и сопоставлять следует средние показатели динамики. Наиболее выразительные, легко воспринимаемые и поддающиеся истолкованию результаты дает сопоставление средних абсолютных приростов, а также средних темпов прироста.

Средние абсолютные приросты можно сравнивать как путем вычитания, так и путем деления. В первом случае следует из среднего прироста за последующий период (\bar{P}_2) вычесть средний прирост за предыдущий период:

$$\Delta \bar{P} = \bar{P}_2 - \bar{P}_1. \quad (9.21)$$

Если сравниваемые приросты положительные, $\Delta \bar{P}$ характеризует либо абсолютное ускорение прироста уровня явления (если $\Delta \bar{P} > 0$), либо абсолютное замедление прироста (если $\Delta \bar{P} < 0$).

Если сравниваемые приросты отрицательны, $\Delta \bar{P}$ характеризует либо абсолютное ускорение снижения уровня (если $\Delta \bar{P} < 0$), либо абсолютное замедление снижения (если $\Delta \bar{P} > 0$).

В тех случаях, когда оба сравниваемых прироста имеют одинаковые знаки, т. е. характеризуют одинаковые по направлению процессы (либо рост, либо снижение уровня), наряду с показателями абсолютного ускорения или замедления ($\Delta \bar{P}$) можно исчислять также коэффициенты ускорения или замедления средней скорости:

$$k_{уск \bar{P}} = \bar{P}_2 : \bar{P}_1, \quad (|\bar{P}_2| > |\bar{P}_1|), \quad (9.22a)$$

$$k_{зам \bar{P}} = \bar{P}_1 : \bar{P}_2, \quad (|\bar{P}_1| > |\bar{P}_2|). \quad (9.22b)$$

Проанализируем, например, следующие данные.

Таблица 9.10

ПОТРЕБЛЕНИЕ ПРОДУКТОВ ПИТАНИЯ В СССР
(на душу населения в год; кг)

	1950 г.	1965 г.	1973 г.
Мясо и сало (включая птицу и субпродукты в натуре)	26	41	53
Молоко и молочные продукты в пересчете на молоко	172	251	307
Сахар	11,6	34,2	40,8
Картофель	241	142	124
Хлебные продукты (хлеб в пересчете на муку, мука, крупа, бобовые, макаронные изделия)	172	156	143

Рассчитаем среднегодовые абсолютные приросты душевого потребления и их абсолютное и относительное изменение в 1966—1973 гг. по сравнению с 1951—1965 гг.

Таблица 9.11

СРЕДНЕГОДОВЫЕ АБСОЛЮТНЫЕ ПРИРОСТЫ ПОТРЕБЛЕНИЯ И ИХ ИЗМЕНЕНИЕ

	$\bar{\Delta}$		$\Delta\bar{\Delta}$	$\kappa\bar{\Delta}$
	1951—1965 гг.	1966—1973 гг.		
Мясо и сало	1,0	1,5	0,5	1,5
Молоко и молочные продукты	5,3	7,0	2,7	1,3
Сахар	1,5	0,8	-0,7	1,9
Картофель	-6,6	-2,2	4,4	3,0
Хлебные продукты	-1,1	-1,6	-0,5	1,5

Таким образом, в 1966—1973 гг. по сравнению с 1951—1965 гг. произошло абсолютное и относительное ускорение прироста потребления мяса и сала — на 0,5 кг или в 1,5 раза, молока и молочных продуктов — на 2,7 кг, или в 1,3 раза. Иначе обстояло дело с потреблением сахара: его прирост замедлился, т. е. уменьшился на 0,7 кг, или в 1,9 раза (1,5:0,8). Потребление картофеля снижалось, однако снижение происходило замедленно: абсолютное замедление снижения составило 4,4 кг, а в относительном выражении снижение замедлилось в 3 раза (6,6:2,2). Напротив, потребление хлебных продуктов снижалось ускоренно: абсолютное ускорение снижения составило 0,5 кг, а в относительном выражении снижение ускорилось почти в 1,5 раза (1,6:1,1).

Средние темпы прироста (или снижения) целесообразно сравнивать только путем деления большего по абсолютной величине на меньший, причем оба сравниваемых темпа должны характери-

зовать одинаковые по направленности процессы, т. е. иметь одинаковые знаки. Получаемые при этом показатели называются коэффициентами ускорения (или замедления) средних темпов прироста:

$$k_{\text{уск}} \bar{\Delta}_{\text{пр}} = \bar{\Delta}_2 : \bar{\Delta}_1 \quad (|\bar{\Delta}_2| > |\bar{\Delta}_1|), \quad (9.23a)$$

$$k_{\text{зам}} \bar{\Delta}_{\text{пр}} = \bar{\Delta}_1 : \bar{\Delta}_2 \quad (|\bar{\Delta}_1| > |\bar{\Delta}_2|). \quad (9.23b)$$

В нашем примере (см. табл. 9.10) среднегодовые темпы прироста (или снижения) потребления составляют (процентов):

	1951—1965 гг.	1966—1973 гг.
Мясо и сало	3,1	3,3
Молоко и молочные продукты	2,3	2,5
Сахар	7,5	2,2
Картофель	-3,5	-1,7
Хлебные продукты	-0,6	-1,1

Таким образом, потребление мяса и сала росло относительно стабильными темпами (коэффициент относительного ускорения средних темпов равен 1,06 (3,3:3,1). Потребление молока и молочных продуктов росло более ускоренными темпами: коэффициент ускорения средних темпов прироста составил 1,09 (2,5:2,3). Потребление сахара тоже росло, но замедленными темпами: они снизились более чем в 3 раза (7,5:2,2). Потребление картофеля снижалось, но в 1966—1973 гг. темпы снижения замедлились примерно вдвое (3,5:1,7). Потребление же хлебных продуктов, снижалось, напротив, ускоренными темпами, которые в 1966—1973 гг. возросли в 1,8 раза.

Рассмотренные выше показатели анализа динамики находят применение в практике любой экономической работы. Они широко используются в директивных указаниях и решениях Коммунистической партии и Советского правительства. Так, в отчетном докладе ЦК КПСС XXIV съезду партии отмечается, что в 1966—1970 гг. «национальный доход страны, использованный на потребление и накопление, увеличивался в среднем за год на 7,1 процента против 5,7 в предыдущей пятилетке. Производительность общественного труда — важнейший показатель эффективности производства — возросла на 37 процентов против 29 в седьмой пятилетке»¹.

4 СПОСОБЫ ГРАФИЧЕСКОГО ИЗОБРАЖЕНИЯ ДИНАМИКИ

Графически динамика явлений наиболее часто изображается в виде столбиковых (см. рис. 4.1) и линейных диаграмм (см. рис. 9.4). Применяются и другие формы диаграмм — фигурные, квадратные, секторные и т. п. (см. рис. 4.6, 5.5). Графики для популяризации чаще строятся в виде столбиковых, а аналитические графики — в виде линейных диаграмм.

¹ Материалы XXIV съезда КПСС, М., Политиздат, 1971, с. 33.

Столбиковые диаграммы более целесообразны, чем линейные, в тех случаях, когда уровни ряда взяты только за отдельные годы с большими или неравными интервалами между ними (на пример, 1940, 1950, 1960, 1970, 1973 гг.), а число уровней невелико (см. рис. 4.1). Порядок следования столбиков должен быть строго хронологическим.

Линейные диаграммы применяются в следующих случаях: а) когда число уровней в ряду динамики велико, б) когда наиболее важным является изображение общей тенденции и характера динамики, в) когда на одном графике необходимо изобразить несколько динамических рядов с целью их сравнения, г) когда более существенным является сопоставление темпов роста, а не уровней и их абсолютных приростов.

Линейные графики практически более удобны для изображения рядов динамики, так как линия лучше выражает непрерывность процесса развития.

Для построения линейной диаграммы используется система прямоугольных координат, на осях которой обычно строится числовая сетка. По оси абсцисс откладывается время, а по оси ординат — либо уровни, либо базисные темпы роста. Каждая точка оси абсцисс выражает определенный момент времени, а отрезки шкалы — периоды (интервалы) времени. Поэтому теоретически моменты времени следует подписывать под точками шкалы, отмеченными черточками, а интервалы (месяцы, годы и т. п.) — под отрезками. Однако практически это теоретическое требование часто не соблюдается, и интервалы также подписываются под точками (см. рис. 9.4, 9.6, 9.8).

Линейная диаграмма (ломаная кривая) наиболее соответствует характеру моментных показателей, уровни которых изображаются вершинами соответствующих ординат. Что касается интервальных показателей, то их уровни теоретически правильно изображать столбиками или ступенчатой кривой (верхние границы столбиков, соединенные вертикальными отрезками), так как они относятся не к моментам, а к периодам времени. Однако из соображений практического удобства динамика интервальных показателей чаще всего изображается в виде ломаной кривой (см. рис. 9.4 и 9.6). Если при этом периоды подписаны под отрезками, то ординаты уровней следует проводить через середины этих отрезков, если же периоды подписаны под точками, то через них проводятся и ординаты уровней (см. рис. 9.3).

При построении линейных диаграмм динамики необходимо строго соблюдать масштаб не только для уровней ряда, но и для времени. Если промежутки времени между датами или периодами неодинаковы, то и ординаты уровней необходимо проводить на соответствующих расстояниях друг от друга. На оси абсцисс в этих случаях обычно подписывают или отмечают черточками только те даты или периоды, уровни которых нанесены на диаграмму (см. рис. 9.4).

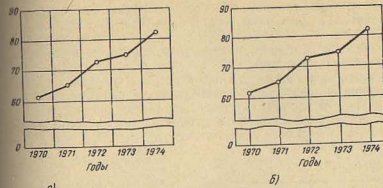


Рис. 9.3. Линейная диаграмма интервального ряда динамики

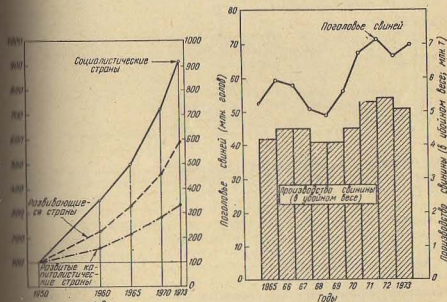


Рис. 9.4. Темпы роста промышленной продукции в социалистических и остальных странах (в процентах к 1950 г.)

Рис. 9.5. Поголовье свиней и производство свинины в СССР

На одном и том же координатном поле можно построить две кривые, уровни которых выражены в различных единицах измерения. В этих случаях по оси ординат устанавливают масштабы отдельно для каждой кривой. В рамках одной сетки кривые можно также комбинировать со столбиками (см. рис. 9.5).

При равномерной шкале на оси ординат прямая линия изображает ряд динамики с равными абсолютными приростами за рав-

ные промежутки времени. При неравных приростах прямая переходит в ломаную, причем чем больше абсолютный прирост, тем больше угол наклона соответствующего отрезка ломаной к оси абсцисс.

Если основной задачей линейного графика является наглядное изображение и сопоставление темпов роста (а не абсолютных приростов), то на оси ординат следует построить логарифмическую шкалу. Тогда прямой линией будет изображаться ряд динамики с равными темпами роста за равные промежутки времени, ломаной — ряд динамики с разными темпами роста. При этом чем больше темп роста, тем больше угол наклона отрезка ломаной к оси абсцисс.

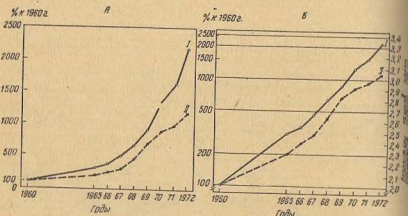


Рис. 9.6. Темпы роста производства средств вычислительной техники (I) и приборов для механизации и автоматизации инженерного и управленческого труда (II) в СССР (А — шкала на оси ординат равномерная, Б — логарифмическая)

На рис. 9.6 темпы роста производства изображены на обычной (равномерной) числовой сетке (А) и на полулогарифмической сетке, где шкала на оси ординат (темпы) — логарифмическая, а на оси абсцисс (время) — равномерная (Б). На полулогарифмической сетке отчетливо видно, что в 1966 г. темп роста средств вычислительной техники (I) несколько замедлился, а затем ускорился, особенно в 1970 г. Производство приборов для механизации и автоматизации инженерного и управленческого труда (II) после 1965 г. росло все ускоряющимися темпами, а после 1970 г. темпы роста снизились. На графике видно также, что в 1968–1969 гг. производство приборов (II) росло быстрее — более высокими темпами, чем производство вычислительной техники (I), тогда как после 1970 г. картина была обратной.

Для изображения изменения структуры могут быть использованы линейные диаграммы. Диаграммы такого рода, называемые слоевыми, могут быть как абсолютными, так и относительными (см. рис. 9.7). Как абсолютные, так и относительные величины

составных частей откладывают на ординатах соответствующих лет последовательно: первая часть (в нашем примере — уголь) — начиная от нуля, вторая — от конца первой и т. д. Так, в 1955 г. удельный вес угля составлял 64,8%, нефти — 21,1%, угля и нефти вместе — 85,9% и т. п.

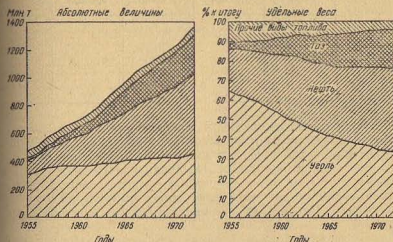


Рис. 9.7. Изменение структуры добычи топлива в СССР (в пересчете на условное топливо)

9. ОСНОВНЫЕ ПРИЕМЫ ОБРАБОТКИ И АНАЛИЗА РЯДОВ ДИНАМИКИ

Смыкание рядов динамики

Иногда уровни явления за одни годы несопоставимы с уровнями за другие годы в связи с территориальными или ведомственными организационными изменениями, изменением методологии исчисления показателя и т. п. Чтобы обеспечить сопоставимость уровней и получить пригодный для анализа единый непрерывный ряд, нужно произвести прямой пересчет уровней, несопоставимых с другими. Однако в распоряжении экономиста иногда нет сведений, необходимых для такого пересчета за все годы, а есть лишь уровни за один и тот же год, исчисленные в старых и в новых границах, по прежней и по новой методологии и т. п. В таких случаях единый динамический ряд может быть получен с помощью особого приема, называемого смыканием рядов динамики.

Допустим, что в 1970 г. произошло изменение границ района. Имеются следующие данные по этому району (гр. 1–2 табл. 9.12).

В связи с изменением границ данные за 1971 и 1972 гг. несопоставимы с данными за 1968–1969 гг. Чтобы сомкнуть эти динамические ряды и получить возможность анализировать динамику среднего удоя за весь период 1968–1972 гг., примем в каждом из них за базу сравнения уровень 1970 г., за который имеют-

Таблица 9.8

СРЕДНИЙ УДОЙ МОЛОКА ОТ ОДНОЙ КОРОВЫ (кг)

Годы	До изменения границ района (в старых границах)	После измене- ния границ (в новых границах)	1970 г. = 100%		Средний удо- й динамики (1970 г. = 100%)
			в старых границах	в новых границах	
А	1	2	3	4	5
1968	1995	...	95	...	95
1969	2058	...	98	...	98
1970	2100	2000	100	100	100
1971	...	2040	...	102	102
1972	...	2100	...	105	105

ся данные как в прежних, так и в новых границах района. В результате получим ряды относительных величин с одинаковой базой сравнения (гр. 3 и 4), которые можно заменить одним суммарным рядом динамики (гр. 5). По данным этого ряда могут быть исчислены темпы роста по отношению к уровню любого года. Могут быть получены и абсолютные уровни за 1968—1969 гг. в новых границах. Так, в 1968 г. средний удой от одной коровы составил 1900 кг ($2000 \times 0,95$).

Следует, конечно, иметь в виду, что результаты, полученные путем подобного смыкания рядов динамики, являются приближенными, т. е. содержат в себе некоторую погрешность.

Характеристика основной тенденции развития (тренда)

Одна из важнейших задач анализа динамики состоит, как уже говорилось выше, в том, чтобы выделить однородные этапы развития явления, а затем вскрыть и охарактеризовать свойственные им закономерности и тенденции. Выделение однородных этапов производится на основе исследования сущности, природы явления и общих законов его развития, т. е. на базе теории той науки, к области изучения которой относится данное явление.

Под основной тенденцией динамики (трендом) понимается общее направление изменения уровня явления, т. е. тенденция к его росту, снижению или стабилизации.

Если уровень явления на изучаемом этапе его развития непрерывно растет или непрерывно снижается, то основная тенденция является явной и отчетливой. Однако тенденция к росту (или снижению) уровня явления может все же при этом носить различный характер в зависимости от того, как с течением времени изменяются ценные абсолютные приросты и ценные темпы прироста, т. е. в зависимости от тенденции изменения ценных показателей динамики. Если, в частности, ценные абсолютные приросты примерно одинаковы, т. е. колеблются около более или менее постоянной величины, то уровень явления имеет тенденцию к росту (снижению) в арифметической прогрессии, т. е. с постоянной абсолютной скоростью. Если же примерно постоянны (ста-

бильны) ценные темпы прироста (снижения), то имеет место тенденция к росту (снижению) уровня явления в геометрической прогрессии, т. е. с постоянной относительной скоростью.

Для количественной характеристики явной и отчетливо выраженной тенденции изменения уровня можно использовать средний абсолютный прирост и средний темп прироста (снижения), исчисляемые по приведенным выше формулам (9.16—9.19), если только ценные показатели динамики не испытывают от года к году очень резких колебаний. Для характеристики же общих результатов действия основной тенденции в этих случаях можно использовать абсолютный прирост и темпы роста и прироста за весь этап развития явления (формулы 9.1, 9.3, 9.5).

Если, однако, ценные показатели динамики, даже оставаясь все время положительными (или отрицательными) резко колеблются от года к году, расчет их средней величины по обычным формулам может дать неправильное представление о средней скорости изменения уровня, соответствующей основной тенденции, т. е. либо преуменьшить, либо преувеличить эту скорость. Соответственно неправильное представление об общих результатах действия основной тенденции может дать в таких случаях и расчет по обычным формулам абсолютного прироста и темпов роста и прироста. К еще большим искажениям и ошибкам при характеристике основной тенденции и результатов ее действия может привести использование приведенных выше обычных формул в тех случаях, когда уровень явления то повышается, то понижается.

Во всех подобных случаях при использовании обычных формул всегда есть риск принять за сравниваемый (y_1) и базисный (y_0) уровни такие, которые не являются достаточно характерными и типичными.

Рассмотрим, например, следующие данные.

Таблица 9.13

ПРОИЗВОДСТВО САХАРА-ПЕСКА В СССР

Годы	Мил. тонн	Годы	Мил. тонн
1961	8,4	1966	9,7
1962	7,8	1967	9,9
1963	6,2	1968	10,8
1964	8,2	1969	10,3
1965	11,0	1970	10,2
1961—1965 (в среднем за год)	8,3	1966—1970 (в среднем за год)	10,2

Если расчет абсолютных приростов, а также темпов роста и прироста за 1966—1970 гг. производить здесь по обычным формулам [(9.1), (9.3), (9.5)], то в качестве базы сравнения следовало

бы принять уровень 1965 г., который значительно превышает не только уровень соседних лет, но и средние уровни за 1961—1965 гг. и 1966—1970 гг. В результате при использовании обычных формул получается, что за 5 лет (1966—1970 гг.) производство сахара-песка уменьшилось на 0,8 млн. т, или почти на 7%. Между тем среднегодовое производство в 1966—1970 гг. по сравнению с 1961—1965 гг. не только не уменьшилось, но, напротив, возросло на 1,9 млн. т, или на 23%.

Безусловно, что среднегодовые уровни, при расчете которых в значительной мере устраняются случайные колебания, более устойчивы и поэтому более характерны и показательны, чем годовые, находящиеся под воздействием кратковременных и преходящих обстоятельств и факторов. Поэтому в тех случаях, когда уровень ряда динамики то повышается, то понижается или же изменяется в одном направлении, но очень неравномерно, так что цепные показатели динамики резко колеблются от года к году, следует сравнивать не годовые, а более типичные и устойчивые среднегодовые уровни.

Расчет показателей анализа динамики и их средних величин в этих случаях следует производить по следующим модифицированным формулам:

$$\bar{P}(\bar{y}) = \bar{y}_{за\ t\ сравниваемых\ лет} - \bar{y}_{за\ t\ базисных\ лет}; \quad (9.24)$$

$$T_{p(\bar{y})} = \frac{\bar{y}_{за\ t\ сравниваемых\ лет}}{\bar{y}_{за\ t\ базисных\ лет}}, \quad (9.25)$$

$$T_{пр(\bar{y})} = \frac{\bar{P}(\bar{y})}{\bar{y}_{за\ t\ базисных\ лет}} = T_{p(\bar{y})} - 1, \quad (9.26)$$

$$\bar{P}(\bar{y}) = \frac{\bar{y}_{за\ t\ сравниваемых\ лет} - \bar{y}_{за\ t\ базисных\ лет}}{t}, \quad (9.27)$$

$$\bar{T}_{p(\bar{y})} = \sqrt[t]{\frac{\bar{y}_{за\ t\ сравниваемых\ лет}}{\bar{y}_{за\ t\ базисных\ лет}}}; \quad (9.28)$$

$$\bar{T}_{пр(\bar{y})} = \bar{T}_{p(\bar{y})} - 1. \quad (9.29)$$

Так, в нашем примере среднегодовой абсолютный прирост и среднегодовой темп роста следует исчислять следующим путем:

$$\bar{P}(\bar{y}) = \frac{\bar{y}_{1966-1970} - \bar{y}_{1961-1965}}{5},$$

$$\bar{T}_{p(\bar{y})} = \sqrt[5]{\frac{\bar{y}_{1966-1970}}{\bar{y}_{1961-1965}}}.$$

Отметим, что именно так были рассчитаны показатели роста производства сельскохозяйственной продукции в седьмой (1961—

1965 гг.), восьмой (1966—1970 гг.) и девятой (1971—1975 гг.) пятилетках¹.

Формулы (9.24) — (9.29) предусматривают сравнение средних уровней за t лет не по длине периоды. Однако в процессе анализа может возникнуть необходимость их сопоставления и за неодинаковые но продолжительности периоды или этапы развития.

Так, например, в соответствии с Директивами XXIV съезда КПСС среднегодовое производство зерна должно возрасти со 167,5 млн. т в 1966—1970 гг. до 195 млн. т в 1971—1975 гг., т. е. на 27,5 млн. т за 5 лет, или на 5,5 млн. т в среднем ежегодно. По истечении трех первых лет девяти пятилетки можно определить среднегодовое производство за эти годы и его абсолютный прирост по сравнению с восьмой пятилеткой: $y_{1971-1973\text{ г.}} - y_{1966-1970\text{ г.}}$. Чтобы исчислить среднегодовое абсолютный прирост, эту величину нужно разделить на половину общей длины сравниваемых периодов, т. е. на 4 года $[(3+5):2]^2$:

$$\bar{P}(\bar{y}) = \frac{\bar{y}_{1971-1973} - \bar{y}_{1966-1970}}{4}.$$

Аналогично можно определить и средний темп роста:

$$\bar{T}_{p(\bar{y})} = \sqrt[4]{\frac{\bar{y}_{1971-1973}}{\bar{y}_{1966-1970}}}.$$

Таким образом, при разной продолжительности базисного и сравниваемого периодов расчет средних показателей динамики следует производить по следующим формулам:

$$\bar{P}(\bar{y}) = \frac{\bar{y}_{за\ t\ сравниваемых\ лет} - \bar{y}_{за\ t\ базисных\ лет}}{0,5(t+s)}, \quad (9.27a)$$

$$\bar{T}_{p(\bar{y})} = \sqrt[\frac{t+s}{2}]{\frac{\bar{y}_{за\ t\ сравниваемых\ лет}}{\bar{y}_{за\ t\ базисных\ лет}}}. \quad (9.28a)$$

Приемы выявления основной тенденции развития (тренда)

Если уровни ряда динамики сильно колеблются, то повышаясь, то понижаясь, общее направление их изменения может быть неясным или выраженным недостаточно отчетливо. Основная тенденция развития оказывается как бы затуманенной постоянными колебаниями уровней, то вверх, то вниз (см., например, динамику производства сахара-песка в табл. 9.13). В таких случаях для выявления основной тенденции (ее направления и характера) необходимо произвести предварительную обработку динамического ряда.

Выявление основной тенденции может представлять собой не только самостоятельную задачу, при решении которой ставится

¹ См., например, Материалы XXIII съезда КПСС, М., Политиздат, 1966, с. 48—49, 245; Материалы XXIV съезда КПСС, М., Политиздат, 1971, с. 34, 133, 155, 261.

² Условно относим каждый средний уровень к центральному году соответствующего периода, т. е. $y_{1966-1970}$ к 1968 г., а $y_{1971-1973}$ к 1972 г., длину периода роста (t) можно найти как разность номеров этих лет: $1972-1968=4$.

Таблица 9.14

ПРОИЗВОДСТВО КАРТОФЕЛЯ В СССР

Годы	Валовой сбор (млн. т)	Годы	Валовой сбор (млн. т)	Годы	Валовой сбор (млн. т)
1959	86,6	1964	93,6	1969	91,8
1960	84,4	1965	88,7	1970	96,8
1961	84,3	1966	87,9	1971	92,7
1962	69,7	1967	95,5	1972	78,3
1963	71,5	1968	102,2	1973	107,7

цель определить общее направление тренда и его характер. Выявление основной тенденции может преследовать и иную цель: выделить в динамическом ряду регулярные, систематические колебания, накладывающиеся на тренд (например, сезонные колебания).

В статистике используются различные приемы обработки ряда динамики с целью выявления направления и характера основной тенденции. Наиболее элементарные из них сводятся либо к простому укрупнению интервалов, либо к расчету средних уровней за такие укрупненные интервалы.

Каждый уровень динамического ряда складывается под воздействием многих различных факторов и обстоятельств. Одни из них действуют на протяжении всего этапа развития с возрастающей либо убывающей силой, формируя основную тенденцию развития (например, совершенствование агротехники, рост производительности труда при производстве сельскохозяйственной продукции и т. п.). Другие факторы и обстоятельства являются временными и преходящими, вызывая отклонения уровня явления от основной тенденции его развития.

В итоговых уровнях за укрупненные интервалы, а также в средних уровнях взаимопоглощаются колебания, вызванные влиянием случайных обстоятельств и факторов, действовавших кратковременно, в результате чего более ясно и отчетливо выступает как общее направление основной тенденции, так и ее характер (ускорение или замедление роста и т. п.).

При использовании элементарных приемов выявления основной тенденции производится в известной мере механически, без каких-либо предварительных допущений о ее характере. Более сложные приемы аналитического выравнивания рядов динамики основаны на использовании методов математической статистики и требуют принятия определенной гипотезы о характере основной тенденции, ее математической форме.

Укрупнение интервалов и их характеристика средними уровнями

Укрупнение интервалов, к которым относятся уровни интервального ряда динамики, заключается в переходе от суточных уровней к недельным или декадным, от декадных — к месячным, от месячных — к кварталным или годовым, от годовых — к многолетним. Если уровни динамического ряда колеблются с более или менее определенной периодичностью (волнообразно), то укрупненный интервал целесообразно взять равным периоду колебаний (длине «полны» цикла). Если же такая периодичность отсутствует, то укрупнение интервалов обычно производят постепенно, переходя от менее крупных к более крупным, пока общее направление тренда не станет ясным и отчетливым. Например, имеются следующие данные (см. табл. 9.14).

При переходе к трехлетним периодам (интервалам) колебания уровней хотя и уменьшаются, но остаются. Четырехлетние уровни (за 1959—1962, 1963—1966, 1967—1970 гг.) увеличиваются не-

прерывно, но уровень за остающиеся 3 года несопоставим с ними. Чтобы получить более или менее сравнимые показатели, можно перейти к среднегодовым уровням, однако и в этом случае тенденция остается недостаточно ясной. Наконец, пятилетние уровни дают такую картину.

Таблица 9.15

УКРУПНЕНИЕ ИНТЕРВАЛОВ И ИХ ХАРАКТЕРИСТИКА СРЕДНИМИ УРОВНЯМИ

Период	Валовой сбор картофеля (млн. т)	
	за весь период	в среднем за год
1959—1963 гг.	396,8	79,4
1964—1968 гг.	467,9	93,6
1969—1973 гг.	467,3	93,5

Добавление к нашему динамическому ряду последующих уровней за 1, 2, 3 или 4 года привело бы, однако, к несопоставимости уровней за более короткий период с пятилетними уровнями. Для получения более или менее сравнимых показателей целесообразно поэтому перейти далее от валовых сборов за пятилетия к среднегодовым сборам за эти периоды (см. последнюю графу табл. 9.15).

Если ряд динамики является моментным, а также в тех случаях, когда уровень ряда представляет собой относительную или среднюю величину, суммирование уровней не имеет смысла, и для выявления основной тенденции укрупненные периоды также следует характеризовать средними уровнями. При укрупнении интервалов число членов динамического ряда сильно сокращается, в результате чего движение уровня внутри укрупненного интервала выпадает из поля зрения. В связи с этим для выявления основной тенденции и более детального изучения ее характера используется также так называемая скользящая средняя.

Сглаживание ряда динамики с помощью скользящей средней заключается в том, что исчисляется средний уровень из определенного числа первых по счету уровней ряда, затем из такого же числа уровней, но начиная со второго по счету, далее — начи-

Таблица 9.16

Сглаживание ряда динамики с помощью
шестилетней скользящей средней и ее центрирование

Шестилетия	Валовой сбор картофеля (млн. т)		Центрированная 6-летняя скользящая средняя	Центральный год 2-х шестилетий
	за 6 лет	в среднем за год (скользящая 6-летняя сред- няя)		
1959—1964	490,4	81,7	$(81,7 + 82,1) : 2 = 81,9$	1962
1960—1965	492,5	82,1	$(82,1 + 82,7) : 2 = 82,4$	1963
1961—1966	496,0	82,7	$(82,7 + 84,5) : 2 = 83,6$	1964
1962—1967	507,2	84,5	$(84,5 + 90,0) : 2 = 87,2$	1965
1963—1968	539,7	90,0	$(90,0 + 93,3) : 2 = 91,6$	1966
1964—1969	559,7	93,3	$(93,3 + 93,8) : 2 = 93,6$	1967
1965—1970	562,9	93,8	$(93,8 + 94,5) : 2 = 94,2$	1968
1966—1971	566,9	94,5	$(94,5 + 92,9) : 2 = 93,7$	1969
1967—1972	557,3	92,9	$(92,9 + 94,9) : 2 = 93,9$	1970
1968—1973	569,5	94,9		

ная с третьего и т. д. Таким образом, при вычислении этим способом средних уровней (звеньев скользящей средней) как бы «скользят» по динамическому ряду от его начала к концу, каждый раз отбрасывая один уровень в начале и добавляя следующий (отсюда название — скользящая средняя).

Каждое звено скользящей средней характеризует средний уровень явления за соответствующий период. При графическом изображении каждое звено принято условно относить к центральному интервалу того периода, за который сделан расчет (для моментного ряда — к центральной дате). Если, например, интервальный ряд динамики за 1960—1974 гг. сглаживается с помощью трехлетней скользящей средней, то первое звено, исчисленное за 1960—1962 гг., следует отнести к 1961 г., а последнее, исчисленное за 1972—1974 гг., — к 1973 г.

Если каждое звено скользящей средней исчисляется из четного числа уровней, то центрального интервала не будет (точнее, центральный интервал не будет совпадать ни с одним исходным календарным интервалом). В таких случаях производится центрирование полученных звеньев скользящей средней путем расчета на их основе звеньев двучленной скользящей средней.

В качестве иллюстрации произведем сглаживание ряда динамики производства картофеля (см. табл. 9.14) с помощью шестилетней скользящей средней. Расчет ее звеньев и центрирование показаны в табл. 9.16.

Первые два звена нецентрированной скользящей средней охватывают 1959—1965 гг., центральным из которых является 1962 г. К этому году и следует отнести первое звено центрированной скользящей средней и т. д.

После сглаживания и центрирования тенденция к росту производства картофеля выступает вполне отчетливо. Более детально, чем при простом укрупнении интервалов, можно проследить и характер этой тенденции. Цепные абсолютные приросты для звеньев центрированной скользящей средней составляют (млн. т):

1963 г.	1964 г.	1965 г.	1966 г.	1967 г.	1968 г.	1969 г.	1970 г.
0,5	1,2	3,6	4,4	2,0	0,6	—0,5	0,2

Таким образом, абсолютная скорость роста была очень неравномерной: сначала рост происходил ускоренно, а затем он стал замедляться. Это наглядно видно и на графике (см. рис. 9.8), где после 1965 г. наклон ломаной, изображающей скользящую среднюю, уменьшается.

Вопрос о том, за какой период следует исчислять каждое звено скользящей средней (т. е. брать ли трех-, четырех-, пятилетнюю и т. д. скользящую среднюю), решается в зависимости от конкретных особенностей динамики. Как и при укрупнении интервалов, если в колебаниях уровня наблюдается определенная периодичность, то период, охватываемый каждым звеном скользящей

средней, следует принять равным периоду колебаний (длине циклической волны) или же кратной ему величине. Так, при наличии месячных уровней, испытывающих ежегодно сезонные колебания, целесообразно использовать 12-месячную (или 24-месячную) скользящую среднюю, а имея дело с квартальными уровнями — четырех- или восьми-квартальную среднюю и т. п. Если же колебания уровней являются беспорядочными, то целесообразно постепенно укрупнять интервал, переходя от двухлетней средней к трехлетней и т. д., пока не выявится отчетливая картина тренда.

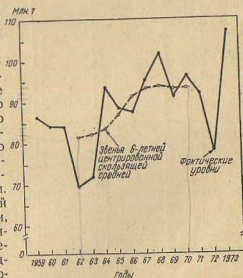


Рис. 9.8. Производство картофеля в СССР

Уменьшение числа звеньев скользящей средней по сравнению с числом исходных уровней ряда динамики, конечно, несколько сужает возможности изучения характера выявленной тенденции в начале и в конце этапа развития. Тем не менее скользящая средняя обладает достаточной *гибкостью*, позволяющей все же уловить изменения характера основной тенденции. Однако скользящая средняя не дает аналитического выражения тренда.

Аналитическое выравнивание

Более сложным приемом выявления основной тенденции развития является аналитическое выравнивание ряда динамики. Оно производится следующим образом.

1. На основе теоретического, качественного анализа сущности и законов развития данного явления выделяется определенный этап его развития и устанавливается характер динамики явления на протяжении этого этапа.

2. Исходя из характера динамики и предположения о той или иной закономерности скорости роста выбирается форма аналитического уравнения, которому графически соответствует определенная линия—прямая, парабола, гипербола, показательная кривая и т. п. Это уравнение аналитически выражает предполагаемую закономерность плавного изменения уровня во времени, т. е. выражает уровень ряда как функцию времени.

Если характер динамики подтверждает предположение о том, что уровень явления растет с более или менее *постоянной абсолютной скоростью*, т. е. с *относительно стабильными ценными абсолютными приростами*, то математическим выражением такой тенденции будет являться *прямая линия*. Аналитическое уравнение прямой применительно к выравниванию таково: $y_t = a_0 + a_1 t$, где y_t — выравниваемые (плавные) уровни, t — время, т. е. порядковый номер интервала или момента времени, a_0 и a_1 — параметры прямой.

Если характер динамики отвечает предположению о росте уровня с более или менее *постоянной относительной скоростью*, т. е. *относительно стабильными ценными темпами прироста*, то в качестве математического выражения этой тенденции следует принять *показательную кривую*: $y_t = a_0 a_1^t$.

В случае, если *ценные абсолютные приросты более или менее равномерно увеличиваются (или уменьшаются)*, т. е. если *уровень растет с равномерно возрастающей (или убывающей) абсолютной скоростью*, в качестве приближенного математического выражения тенденции можно принять *параболу второго порядка*: $y_t = a_0 + a_1 t + a_2 t^2$.

Выбор формы аналитического уравнения, т. е. вида кривой¹, обычно является несколько условным, так как процесс развития любого общественного явления, конечно, не укладывается строго ни в одну математическую формулу. Плавное изменение уровня лишь приблизительно выражает тенденцию развития и только удов-

но может рассматриваться как функция времени. В действительности закономерное изменение уровня во времени обусловлено действием целого комплекса социально-экономических условий и факторов.

3. После выбора вида кривой вычисляются ее параметры¹. Расчет параметров производится, как правило, методом наименьших квадратов. Это означает, что ставится задача: из бесконечного числа кривых выбранного вида найти ту, которая обращает в минимум сумму квадратов отклонений фактических уровней динамического ряда от соответствующих им во времени выровненных уровней, т. е. уровней, лежащих на искомой кривой. Математически это условие выражается так:

$$\sum (y - \hat{y}_t)^2 = \min,$$

где y — фактические уровни ряда динамики,

\hat{y}_t — соответствующие им во времени выровненные уровни, расположенные на искомой прямой или кривой.

Параметры кривой, удовлетворяющей этому условию, могут быть найдены путем решения системы нормальных уравнений.

Если фактические уровни динамического ряда изобразить графически в виде ломаной линии (см. рис. 9.9), то искомая кривая, изображающая плавное изменение уровня, должна проходить к вершинам ломаной в среднем ближе, чем любая другая кривая данного вида.

4. На основе найденного уравнения кривой рассчитываются выровненные уровни, соответствующие во времени фактическим уровням ряда динамики.

Таким образом, техникой выравнивания заключается в замене фактических уровней такими плавно изменяющимися уровнями, которые в среднем менее всего отклонялись бы от фактических² и имели бы определенное аналитическое выражение, соответствующее общему направлению и характеру тренда.

Хотя, как уже отмечалось, выравнивание рядов динамики и содержит в себе некоторые условности, однако в ряде случаев оно является полезным техническим приемом, облегчающим выявление основной тенденции и изучение ее характера. Помогает оно изучать и другие особенности динамики, в частности сезонные колебания (см. ниже).

Рассмотрим технику выравнивания ряда динамики по прямой:

$$\hat{y}_t = a_0 + a_1 t.$$

Параметры a_0 и a_1 искомой прямой, удовлетворяющие принципу наименьших квадратов, находятся путем решения следующей системы нормальных уравнений:

$$\begin{aligned} a_0 n + a_1 \sum t &= \sum y, \\ a_0 \sum t + a_1 \sum t^2 &= \sum y t, \end{aligned}$$

¹ См. список на с. 328.

² В смысле наименьшей средней квадратической величины отклонений.

¹ Прямую можно рассматривать как частный случай кривой.

где y — фактические (эмпирические) уровни ряда динамики;
 n — число уровней;

t — время (порядковый номер интервала или момента времени).

Расчет параметров значительно упрощается, если за начало отсчета времени принять центральный интервал (или момент) рассматриваемого этапа. При нечетном числе уровней получим тогда такие значения t :

Годы	1966	1967	1968	1969	1970	1971	1972
t	-3	-2	-1	0	1	2	3

Если же количество уровней в выравниваемом ряду четное, то значения t устанавливаются так, как это сделано в гр. 2 табл. 9.17. Это равнозначно измерению времени не в годах, а в полугодиях.

В обоих случаях $\Sigma t = 0$, в результате чего система уравнений принимает вид:

$$a_0 n = \Sigma y,$$

$$a_1 \Sigma t^2 = \Sigma yt,$$

откуда

$$a_0 = \Sigma y : n, \quad (9.30a)$$

$$a_1 = \Sigma yt : \Sigma t^2. \quad (9.30б)$$

Проиллюстрируем выравнивание ряда динамики по прямой на следующем примере.

Таблица 9.17

РАСЧЕТНАЯ ТАБЛИЦА ПРИ ВЫРАВНИВАНИИ ПО ПРЯМОЙ
 РЯДА ДИНАМИКИ ВАЛОВОГО СБОРА ЗЕРНА В СССР

Годы	Валовой сбор зерна (млн. т.) y	Обозначение времени t	yt	t^2	Выравненные уровни \hat{y}_t
A	1	2	3	4	5
1964	152	-9	-1 368	81	137,6
1965	121	-7	-847	49	144,4
1966	171	-5	-855	25	151,2
1967	148	-3	-444	9	158,0
1968	170	-1	-170	1	164,8
1969	162	1	162	1	171,6
1970	187	3	561	9	178,4
1971	181	5	905	25	185,2
1972	168	7	1 176	49	192,0
1973	222	9	1 998	81	198,8
Итого	1 682	0	1 118	330	1 682,0

Валовой сбор зерна имеет явную тенденцию к росту (см. рис. 9.9), причем положительные и отрицательные цепные абсолютные

приросты чередуются. Однако уже простое укрупнение интервалов до двух лет и расчет среднегодовых абсолютных приростов по формуле (9.24) показывает, что они относительно стабильны: $(Y_{1966/67} - Y_{1964/65}) : 2 = 11,0$; $(Y_{1968/69} - Y_{1966/67}) : 2 = 3,2$; $(Y_{1970/71} - Y_{1968/69}) : 2 = 9,0$; $(Y_{1972/73} - Y_{1970/71}) : 2 = 5,5$. Поэтому динамический ряд приближенно можно выравнить по прямой.

Подставляя в формулы (9.30 а, б) итоговые суммы из табл. 9.17, получим:

$$a_0 = \frac{\Sigma y}{n} = \frac{1682}{10} = 168,2, \quad a_1 = \frac{\Sigma yt}{\Sigma t^2} = \frac{1118}{330} = 3,4.$$

Отсюда уравнение иско-
 мой прямой будет: $\hat{y}_t =$
 $= 168,2 + 3,4t$.

Путем подстановки в это уравнение соответствующих значений t (см. гр. 2 табл. 9.17) найдем выравненные уровни \hat{y}_t .

Так, для 1964 г. ($t = -9$) получим: $\hat{y}_{64} = 168,2 + (3,4 \times (-9)) = 137,6$ и т. д. (см. гр. 5 табл. 9.17).

На рис. 9.9 изображены фактический (эмпирический) ряд динамики и выравненные его уровни (выравненный динамический ряд).

Отметим, что аналитическое выравнивание ряда динамики позволяет не только сделать четким общее направление основной тенденции, но одновременно дает также численную характеристику тренда в виде различных показателей средней скорости изменения уровня. В частности, при выравнивании по прямой параметр a_1 представляет собой абсолютный прирост выравненного уровня за единицу времени t .

При нечетном числе уровней, когда соседние значения отличаются одно от другого на 1, параметр a_1 выражает абсолютный прирост за интервал времени, к которому относятся уровни интервального ряда, или же за интервал между датами моментного ряда динамики, т. е. ежегодный, ежемесячный и т. п. прирост. Если же, как в нашем примере, число уровней четное и соседние значения t отличаются на 2, то счет времени производится уже половинными исходных интервалов (в нашем примере — полугодиями). Соответственно и a_1 характеризует прирост за половину исходного интервала.

В нашем примере $a_1 = 3,4$ означает, что прирост валового сбора в расчете на полугодие составлял 3,4 млн. т, т. е. ежегодно валовой сбор возрастал на 6,8 млн. т.

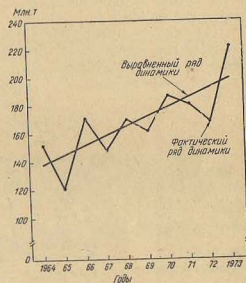


Рис. 9.9. Валовой сбор зерна в СССР

При выравнивании ряда динамики по показательной кривой $\hat{y}_t = a_0 a_1^t$ параметр a_1 характеризует темп роста выравненного уровня за единицу времени t .

Показательная кривая путем логарифмирования приводится к линейному виду относительно $\lg \hat{y}_t$, $\lg a_0$ и $\lg a_1$:

$$\lg \hat{y}_t = \lg a_0 + t \lg a_1.$$

Поэтому техника выравнивания по показательной кривой аналогична технике выравнивания по прямой с той лишь разницей, что при этом сначала рассчитываются $\lg a_0$, $\lg a_1$ и $\lg \hat{y}_t$, а затем путем потенцирования определяются параметры a_0 и a_1 , а также выравненные уровни \hat{y}_t .

При выравнивании ряда динамики по параболе второго порядка ($\hat{y}_t = a_0 + a_1 t + a_2 t^2$) ее параметры (при $\Sigma t = 0$) рассчитываются следующим образом:

$$\left. \begin{aligned} a_1 &= \frac{\Sigma y t}{\Sigma t^2}, \\ a_0 n + a_2 \Sigma t^2 &= \Sigma y, \\ a_0 \Sigma t^2 + a_2 \Sigma t^4 &= \Sigma y t^2. \end{aligned} \right\} \quad (9.31a, б, в)$$

Абсолютная скорость роста выравненного уровня в этом случае с течением времени изменяется (она равна $a_1 + 2a_2 t$), причем за единицу времени t скорость увеличивается (или уменьшается) на $2a_2$. Таким образом, параметр a_2 представляет собой половину прироста скорости за единицу времени, т. е. половину ускорения (если соседние значения отличаются на 2, то ускорение равно $8a_2$).

Некоторые общественные явления и процессы подвержены сезонным колебаниям, т. е. их уровень из года в год в определенные месяцы повышается, а в другие — снижается. Эти *внутригодовые колебания*, имеющие более или менее регулярный, устойчивый характер, и называются *сезонными колебаниями*.

Сезонные колебания могут быть обусловлены сменой времен года, а также различными социально-экономическими факторами. Особенно наглядно сезонные колебания проявляются в сельском хозяйстве, где, например, получение большинства продуктов растениеводства происходит только в определенные месяцы. Это приводит в свою очередь к сезонным колебаниям в перевозках и в работе промышленных предприятий, перерабатывающих сельскохозяйственное сырье. Наблюдаются сезонные колебания и в других отраслях народного хозяйства — в строительстве, в торговле и т. д.

Сезонные подъемы и спады часто приводят к разного рода нежелательным результатам. Они вызывают, в частности, неравномерное в течение года использование трудовых ресурсов и производственных мощностей, что нередко ведет к снижению произво-

дительности труда, повышению себестоимости продукции, ухудшению обслуживания населения и т. п. В связи с этим, как правило, стремятся либо ликвидировать, либо уменьшить, смягчить сезонные колебания.

Отсюда вытекает необходимость изучения сезонности и измерения сезонных колебаний. Это измерение позволяет судить об эффективности и результатах мер по изжитию или смягчению сезонности. Вместе с тем измерение сезонных колебаний необходимо также для того, чтобы при внутригодовом планировании предвидеть и учесть такие сезонные колебания, которые пока еще не удастся устранить. Механическая разбивка годового плана по месяцам и кварталам могла бы привести при этом к серьезным неувязкам.

Существуют различные способы выявления и числовой характеристики сезонных колебаний. Обычно для их измерения используются специальные показатели — индексы сезонности. *Индекс сезонности* — это процентное отношение фактического уровня явления за тот или иной месяц (или квартал) к выравненному уровню за тот же месяц (или квартал):

$$I_{\text{сез}} = \frac{y_{\text{факт}}}{y_{\text{выравни}}} \times 100. \quad (9.32)$$

Поскольку в каждом году сезонные колебания обычно имеют свои особенности, индексы сезонности, как правило, исчисляют не за один год, а за несколько лет (обычно за 3—4 года). Чтобы получить устойчивые индексы, свободные от случайных особенностей отдельных лет и отражающие типичный характер сезонных колебаний, из индексов за одноименные месяцы затем вычисляют простые средние арифметические:

$$\bar{I}_{\text{сез}} = \Sigma I_{\text{сез}} : n. \quad (9.33)$$

В зависимости от наличия или отсутствия в динамическом ряду отчетливо выраженной тенденции к росту или снижению выравненный уровень в формуле (9.32) исчисляется по-разному.

Если в динамическом ряду нет ясно выраженной тенденции к росту или снижению, в качестве выравненного уровня можно принять *среднюю арифметическую* из месячных уровней за весь изучаемый период. Так, если взят трехлетний период, в качестве *выравни* берется средняя из 36 месячных уровней. В этом случае в числителе индекса также можно брать средние уровни за одноименные месяцы, т. е. при трехлетнем периоде — среднюю из трех январских уровней, затем — из трех февральских уровней и т. д. Это сразу дает средние индексы сезонности, так что осреднение с помощью формулы (9.33) уже не требуется.

Если же уровень явления из года в год существенно повышается или понижается, метод простой средней арифметической для измерения сезонных колебаний не пригоден, так как колебания около постоянного среднего уровня дадут искаженную кар-

тину сезонной волны. Поэтому при наличии отчетливой и достаточно сильной тенденции к росту или снижению уровня от года к году следует предварительно выделить тренд либо путем сглаживания ряда динамики с помощью 12-месячной скользящей средней, либо путем аналитического выравнивания. Это позволит измерить колебания около уровней, выражающих общую тенденцию. В качестве таких выровненных уровней в формуле (9.32) и можно использовать либо звенья 12-месячной центрированной скользящей средней, либо уровни, полученные при аналитическом выравнивании.

Рассчитаем индексы сезонности по следующим данным одного колхоза.

Таблица 9.18
РАСЧЕТ ИНДЕКСОВ СЕЗОННОСТИ ОБЪЕМА ТРАКТОРНЫХ РАБОТ В КОЛХОЗЕ

Месяц	Объем тракторных работ (тыс. га мягкой пахоты)			Центрированная 12-месе- чная скользящая средняя			Индексы сезонности (в процентах)			В среднем за 1970— 1972 гг.
	1970 г.	1971 г.	1972 г.	1970 г.	1971 г.	1972 г.	1970 г.	1971 г.	1972 г.	
A	1	2	3	4	5	6	7	8	9	10
I	1,1	1,3	1,6	—	2,37	2,65	—	55	60	58
II	1,1	1,3	1,7	—	2,40	2,67	—	54	64	59
III	1,3	1,6	1,8	—	2,42	2,69	—	66	67	66
IV	2,1	2,4	2,7	—	2,45	2,72	—	98	99	98
V	3,4	3,8	4,0	—	2,47	2,75	—	154	145	150
VI	2,4	2,7	3,0	—	2,48	2,78	—	109	108	108
VII	2,3	2,6	2,9	2,22	2,50	—	104	104	—	104
VIII	3,3	3,6	3,8	2,24	2,53	—	147	142	—	144
IX	3,6	4,0	4,4	2,26	2,56	—	159	156	—	158
X	3,0	3,2	3,6	2,29	2,58	—	131	124	—	128
XI	1,7	1,9	2,2	2,32	2,60	—	73	73	—	73
XII	1,3	1,5	1,8	2,35	2,62	—	55	57	—	56
Итого	26,6	29,9	33,5	—	—	—	—	—	—	1202

Как видно из годовых итогов, объем работ из года в год существенно увеличивается. За два года он возрос почти на 7 тыс. га, или на 26%. Чтобы получить достаточно точную картину сезонных колебаний, выделим тренд с помощью 12-месячной скользящей средней.

Первая 12-месячная сумма равна 26,6, а вторая—26,8. Средняя из этих величин (26,7), после деления на 12, дает первое звено скользящей средней, центрированной на седьмом месяце: $26,7 : 12 = 2,22$. Эту величину отнесем к июлю 1970 г. и т. д.

Вычислив все звенья скользящей средней (см. гр. 4—6 табл. 9.18), найдем далее индексы сезонности. Для этого фактический уровень каждого месяца, начиная с июля 1970 г., нужно разделить на соответствующее звено скользящей средней. Так, для июля 1970 г. получим: $(2,3 : 2,22) \times 100 = 104\%$, для августа

1970 г.— $(3,3 : 2,24) \times 100 = 147\%$ и т. д. (см. гр. 7—9 табл. 9.18).

За каждый месяц мы получим два индекса сезонности. Простая средняя арифметическая из них даст средние индексы за изучаемый период (см. гр. 10 табл. 9.18).

Так как средний индекс сезонности для всех 12 месяцев должен быть равен 100%, то сумма индексов должна составить 1200. Если она не равна этой величине, необходимо сделать корректировку, умножив индексы всех месяцев на отношение $(1200 : \Sigma I)$. В нашем примере это отношение очень близко к 1, а средние индексы вычислялись с округлением, так что корректировка практически не нужна.

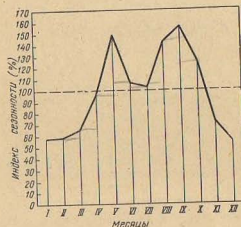


Рис. 9.10. Сезонные колебания объема тракторных работ в колхозе (средние индексы сезонности за 1970—1972 гг.)

Размах и характер сезонных колебаний особенно наглядны при графическом изображении (см. рис. 9.10). Еще одним вариантом графиков сезонности являются радиальные диаграммы, которые строятся в полярной системе координат. Окружность делится на 12 равных частей и из центра через точки деления проводятся лучи. Каждый из них обозначает определенный месяц. Величина радиуса принимается за среднемесячный уровень или за 100%, и в соответствии с этим масштабом на лучах, начиная от центра окружности, откладываются отрезки, изображающие месячные уровни или индексы сезонности. Концы этих отрезков последовательно соединяются друг с другом и образуется замкнутая фигура — двенадцатиугольник (см. рис. 9.11).

В качестве обобщающих показателей, позволяющих судить о том, уменьшились или углубились сезонные колебания, могут

быть использованы либо среднее линейное, либо среднее квадратическое отклонение индексов сезонности от 100%:

$$\sigma = \sqrt{\frac{\sum (I_{сез} - 100)^2}{12}} \quad (9.34)$$

Эти же показатели можно использовать для сравнения сезонных колебаний разных явлений и т. п.

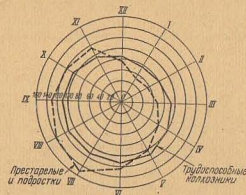


Рис. 9.11. Сезонные колебания численности работников колхозников (по данным выборочного обследования за 1967 г.; среднемесячный уровень = 100%)

Сравнительный анализ нескольких рядов динамики

Важное место при изучении развития социально-экономических явлений занимает сравнительный анализ нескольких рядов динамики. При этом может ставиться задача сопоставления:

- динамики одного и того же явления на разных территориях,
- динамики явления в целом и его составных частей,
- динамики различных явлений, между которыми существуют либо функциональная, либо корреляционная связь (например, общее производство продукции, численность населения и производство продукции на душу населения или себестоимость центнера зерна и урожайность зерновых и т. п.),
- динамики различных явлений, характеризующих одну сторону или область производства, обслуживания и т. п. (например, производство телевизоров, бытовых холодильников, мотоциклов и других товаров культурно-бытового назначения).

При сравнительном анализе динамики одноименных явлений, относящихся к различным территориям или являющихся составными частями целого, можно сопоставлять не только направление и характер основных тенденций и относительные показатели динамики (темпы развития), но и абсолютные величины: уровни, абсолютные приросты, абсолютные ускорения,

Сравним, например, добычу нефти в 1960—1972 гг. в западных и восточных¹ районах СССР.

Таблица 9.19

ДОБЫЧА НЕФТИ В ЗАПАДНЫХ И ВОСТОЧНЫХ РАЙОНАХ СССР

Годы	Добыча (млн. т)			Добыча в восточных районах в процентах	
	всего	в том числе в районах		к общей добыче	к добыче в западных районах
		западных	восточных		
1960	148	134	14	9	10
1965	243	213	30	12	14
1972	400	270	130	32	48

В 1960 г. добыча нефти в восточных районах составляла менее одной десятой части общей добычи и была почти в 10 раз меньше, чем добыча в западных районах страны. За пять лет (1961—1965 гг.) добыча в восточных районах возросла на 16 млн. т, тогда как в западных районах прирост составил 79 млн. т. Однако по темпам роста восточные районы опережали западные: на востоке страны добыча нефти за пятилетие более чем удвоилась, а в западных районах увеличилась только в 1,6 раза. В результате в 1965 г. добыча нефти в восточных районах составила уже 12% общесоюзной добычи.

Особенно резко возросла добыча нефти на востоке в следующие 7 лет: она увеличилась на 100 млн. т, или в 4,3 раза, тогда как в западных районах добыча возросла только на 57 млн. т, или менее чем в 1,3 раза. Таким образом, восточные районы не только догнали, но и перегнали западные по абсолютному размеру прироста добычи: они дали почти две трети общесоюзного прироста (100 из 157 млн. т.).

Расчет среднегодовых абсолютных приростов и темпов прироста дает следующие результаты:

Таблица 9.20

	1961—1965 гг.	1966—1972 гг.
Среднегодовой абсолютный прирост (млн. т):		
в западных районах	15,8	8,1
в восточных районах	3,2	14,3
Среднегодовой темп прироста (процент):		
в западных районах	9,7	3,4
в восточных районах	16,5	23,3

¹ Восточные районы: Уральский, Западно-Сибирский, Восточно-Сибирский, Дальневосточный, Среднеазиатский, Казахстанский.

Следовательно, в 1966—1972 гг. по сравнению с 1961—1965 гг. в восточных районах произошло ускорение абсолютного и относительного прироста: среднегодовой абсолютный прирост увеличился с 3,2 до 14,3 млн. т, т. е. в 4,5 раза, а среднегодовой темп прироста — с 16,5 до 23,3%, т. е. почти в 1,5 раза. В западных же районах ежегодный прирост сильно замедлился: ежегодный абсолютный прирост уменьшился с 15,8 до 8,1 млн. т, т. е. почти вдвое, а ежегодный темп прироста — с 9,7 до 3,4%, т. е. почти втрое. В результате разрыв в уровне добычи нефти на западе и востоке страны существенно сократился: в 1972 г. восточные районы дали почти половину добычи в западных районах или почти одну треть общесоюзной добычи.

При сравнительном анализе рядов динамики для большей наглядности направления развития и его результатов сопоставляемые ряды целесообразно приводить к одному основанию, т. е. к общей базе сравнения, принимаемой за 1 или за 100%. В зависимости от цели анализа в качестве общей базы (основания) каждого ряда могут быть приняты: а) начальный уровень, б) какой-либо другой характерный уровень, в) средний уровень за тот или иной период (в том числе за весь изучаемый период).

Если уровни сравниваемых рядов систематически растут (или снижаются), за базу сравнения удобно принять начальный уровень. Если же уровни то повышаются, то понижаются, базу сравнения целесообразно расширить, приняв за нее средний уровень. Это сделает базу сравнения более характерной, типичной и устойчивой. В частности, при отсутствии явной тенденции к росту или снижению, а также при волнообразных, периодических колебаниях уровней в качестве общей базы сравнения целесообразно принять средний уровень за весь период.

Приведем к одному основанию ряды динамики добычи нефти в западных и восточных районах СССР, приняв за общую базу сравнения уровень 1960 г.

Таблица 9.21

РОСТ ДОБЫЧИ НЕФТИ В ЗАПАДНЫХ И ВОСТОЧНЫХ РАЙОНАХ СССР
(в процентах к 1960 г.)

Годы	Общая добыча	в том числе	
		в западных районах	в восточных районах
1960	100	100	100
1965	164	159	214
1972	270	202	929

Как видно из данных этой и предыдущей таблиц, добыча нефти на востоке страны росла значительно быстрее, чем по стране в целом и особенно чем в западных районах. Иными словами, по темпам роста добычи нефти в 1961—1972 гг. восточные районы страны опережали западные ее районы.

Если сопоставляется динамика одного и того же (одноименного) явления на разных территориях или динамика составных частей целого, для числовой характеристики процесса опережения можно сравнивать любые показатели динамики: цепные, средние и базисные абсолютные приросты, темпы роста и темпы прироста. При этом сравниваемые показатели должны быть исчислены за один и тот же период.

Абсолютные приросты (P) можно сравнивать либо путем вычитания из большего прироста меньшего ($P_A - P_B$), либо путем деления большего прироста на меньший ($P_A : P_B$). В последнем случае оба прироста должны быть либо положительными, либо отрицательными.

Коэффициент, показывающий, во сколько раз один абсолютный прирост больше другого, может быть назван коэффициентом опережения по абсолютному приросту, или коэффициентом абсолютного опережения:

$$k_{\text{абс. опережения}} = P_A : P_B \quad (P_A > P_B), \quad (9.35)$$

где A и B — сравниваемые части целого или сравниваемые территории.

При сравнении средних абсолютных приростов в формуле (9.35) вместо P берется \bar{P} , а коэффициент абсолютного опережения также будет средним (среднегодовым и т. п.).

Так, в нашем примере (см. табл. 9.20) среднегодовой абсолютный прирост добычи нефти в 1961—1965 гг. составлял: в западных районах — 15,8 млн. т, а в восточных — 3,2 млн. т. Следовательно, в западных районах средняя абсолютная скорость роста добычи была в 5 раз выше. В 1966—1972 гг. средний прирост в западных районах составил 8,1 млн. т, а в восточных — 14,3 млн. т, т. е. в 1,8 раза больше. Таким образом, если в 1961—1965 гг. по средней абсолютной скорости роста добычи западные районы опережали восточные в 5 раз, то в 1966—1972 гг., наоборот, восточные районы опережали западные в 1,8 раза.

Темпы развития целесообразно сравнивать только путем деления большего из них на меньший. При этом оба сравниваемых темпа должны характеризовать одинаковый по направлению процесс, т. е. либо рост, либо снижение уровня динамического ряда.

Коэффициент, показывающий во сколько раз один темп больше другого, называется коэффициентом опережения по темпам (роста или прироста), или, короче, коэффициентом относительного опережения:

$$k_{\text{отн. опережения}} (T_p) = \frac{T_p(A)}{T_p(B)} \quad (T_A > T_B). \quad (9.36)$$

Если сравниваются темпы прироста или среднегодовые темпы роста или прироста, в формуле (9.36) вместо T_p берутся соответственно $T_{\text{пр}}$, T_p или $\bar{T}_{\text{пр}}$. При сравнении среднегодовых темпов коэффициент относительного опережения также будет среднегодовым.

Так, в нашем примере (см. табл. 9.20) по средней относительной скорости роста добычи нефти восточные районы значительно опережали западные, причем среднегодовой коэффициент опережения увеличился с 1,7 в 1961—1965 гг. (16,5:9,7) до 6,7 в 1966—1972 гг. (23,3:3,4). В целом за 1961—1972 гг. среднегодовые темпы прироста составили: в западных районах—6%, в восточных—20,4%; коэффициент опережения равен 3,4. Иными словами, в 1961—1972 гг. добыча нефти в восточных районах страны в относительном выражении росла в 3,4 раза быстрее, чем в западных.

Если анализируется динамика различных явлений, то можно сопоставлять только направление и характер их основных тенденций и относительные показатели динамики (темпы развития и коэффициенты их ускорения или замедления). Аналогичное положение имеет место в тех случаях, когда анализируется динамика односторонних явлений в разных странах, причем абсолютные уровни явления в этих странах несопоставимы вследствие различной методики расчета показателей и различной их денежной оценки. Так обстоит дело, например, при сравнительном анализе динамики общего объема промышленной или сельскохозяйственной продукции, производительности труда, национального дохода и т. п. в разных странах.

Ряды динамики, уровни которых непосредственно несопоставимы, обычно приводятся к одному основанию. Это позволяет выявить различия или сходство в направлении и характере основных тенденций и охарактеризовать затем эти тенденции с помощью средних темпов развития, коэффициентов ускорения или замедления. Следует иметь в виду, что для рядов динамики, приведенных к одному основанию, могут быть исчислены и сопоставлены не только коэффициенты ускорения темпов роста или прироста, но и коэффициенты ускорения абсолютных приростов. Для таких рядов динамики могут быть исчислены также коэффициенты относительного опережения.

Рассмотрим, например, следующие данные.

Таблица 9.22

ТЕМПЫ РОСТА ПРОИЗВОДИТЕЛЬНОСТИ ТРУДА В ПРОМЫШЛЕННОСТИ
СССР и США

(в расчете на одного работающего; в процентах к 1960 г.)

Годы	СССР	США
1960	100	100
1965	125	126
1972	185	155

Абсолютные уровни производительности труда здесь неизвестны. Однако можно найти абсолютные приросты, в том числе среднегодовые, измеряя их в так называемых пунктах роста, т. е.

в сотых частях (процентах) базисного уровня. Так, в СССР абсолютный прирост за 1961—1965 гг. составил 25 пунктов (125—100), или 5,0 пункта в среднем ежегодно. В 1966—1972 гг. абсолютный прирост составил 60 пунктов (185—125), или 8,6 пункта в среднем ежегодно. Следовательно, среднегодовой абсолютный прирост в СССР ускорился в 1,7 раза ($\sqrt[7]{8,6:5,0} = 1,7$). Напротив, в США среднегодовой абсолютный прирост снизился с 5,2 пункта в 1961—1965 гг. до 4,1 пункта в последние 7 лет. Коэффициент замедления составил 1,3 (5,2:4,1).

Расчет среднегодовых темпов прироста дает следующие результаты:

	СССР	США
1961—1965 гг.	4,6	4,7
1966—1972 гг.	5,8	3,0
1961—1972 гг.	5,3	3,7

Таким образом, в СССР средние темпы прироста были стабильны и даже несколько возросли, а в США снизились в 1,5 раза. Если в 1961—1965 гг. средние темпы прироста в СССР и США были примерно одинаковы, то в 1966—1972 гг. СССР опережал США в 1,9 раза (5,8:3,0). В целом в 1961—1972 гг. СССР опережал США по среднегодовым темпам прироста производительности труда в 1,4 раза ($\sqrt[7]{5,3:3,7} = 1,4$).

При сравнительном анализе динамики часто строят ряды динамики производных показателей: относительных и средних величин. Так, сопоставляя динамику производства зерна и посевных площадей зерновых культур, можно получить динамический ряд средней урожайности зерновых; путем сопоставления динамики общего производства молока и динамики среднего поголовья коров можно построить динамический ряд среднего удоя молока от одной коровы и т. п.

Следует иметь в виду, что если между тремя показателями существует функциональная зависимость вида $y = xz$, то такая же зависимость существует между темпами роста этих показателей. Следовательно, на основе темпов роста любых двух из этих показателей можно найти темп роста третьего показателя¹.

Так, в 1972 г. по сравнению с 1950 г. городской жилищный фонд СССР (общая, полезная площадь жилищ) увеличился в 3,24 раза, а численность городского населения—в 2,00 раза. Обозначив жилищный фонд через F , численность населения через N , а обеспеченность населения полезной площадью через $O = F:N$, найдем, что темп роста обеспеченности населения полезной площадью равен темпу роста жилищного фонда, деленному на темп роста численности населения. В нашем примере темп роста обеспеченности составляет 1,62 (3,24:2,00), т. е. полезная площадь жилищ, приходящаяся на одного городского жителя, возросла на 62%.

¹ Подробное этог вопрос рассматривается в гл. X.

Если теоретический, социально-экономический анализ показывает, что между уровнями двух (или нескольких) рядов динамики может существовать корреляционная связь, то возникает задача ее измерения. Однако применение методов корреляции и регрессии к рядам динамики имеет ряд особенностей ввиду возможной зависимости между уровнями каждого динамического ряда. Поэтому коррелируются не сами уровни динамических рядов, а либо их отклонения от трендов ($y - \bar{y}_t$), либо цепные абсолютные приросты (первые разности: $y_t - y_{t-1}$). Подробное рассмотрение этого вопроса выходит, однако, за рамки программы данного курса.

Интерполяция и экстраполяция рядов динамики

Иногда один или несколько промежуточных уровней ряда динамики отсутствуют и возникает необходимость приближенного расчета их величин на основе известных уровней ряда. Такой расчет называется интерполяцией.

Интерполяция производится исходя из предположения о той или иной закономерности изменения уровней ряда за рассматриваемый период. Поэтому точность результатов интерполяции зависит, конечно, от того, насколько близкой к действительности является предполагаемая закономерность. Точность результатов зависит также от устойчивости этой закономерности, от стабильности тех или иных показателей динамики, т. е. от степени колеблемости уровней ряда около предполагаемого тренда. Это обуславливается в свою очередь большей или меньшей однородностью условий развития явления.

Чаще всего при интерполяции исходят из предположения о равномерном изменении уровня, т. е. о стабильности либо абсолютных приростов, либо темпов роста. В соответствии с этим интерполяция производится либо на основе среднего абсолютного прироста (\bar{P}), либо на основе среднего темпа роста (\bar{T}_y). В свою очередь \bar{P} и \bar{T}_y , используемые для интерполяции, могут быть исчислены либо на основе только двух уровней, соседних с неизвестным (или неизвестными), либо на основе большего числа уровней ряда.

Допустим, что известны следующие данные (отсутствующий в таблице уровень 1967 г. в действительности составляет 433,0 тыс. км).

Таблица 9.23

ПРОТЯЖЕННОСТЬ АВТОМОБИЛЬНЫХ ДОРОГ С ТВЕРДЫМ ПОКРЫТИЕМ В СССР
(на конец года; тыс. км)

	Годы						
	1964	1965	1966	1967	1968	1969	1970
Протяженность дорог	351,7	378,3	405,6	...	456,4	483,2	511,6
Абсолютный прирост за год	...	26,6	27,3	26,8	28,4

Как видно из данных таблицы, годовые абсолютные приросты относительно стабильны.

Полагая, что прирост в 1967 г. был таким же, как и в 1968 г., и был равен, следовательно, среднему годовому приросту за эти два года $[(456,4 - 405,6) : 2 = 25,4]$, найдем «неизвестный» уровень 1967 г.:

$$y_{1967} = y_{1966} + \bar{P}_{1967-1968} = 405,6 + 25,4 = 431,0 \text{ (тыс. км)}.$$

Это лишь на 2 тыс. км, или менее чем на 0,5%, отличается от действительного уровня 1967 г. (433,0 тыс. км).

Тот же результат можно получить, исчислив простую среднюю арифметическую из уровней 1966 и 1968 гг. $(405,6 + 456,4) : 2 = 431,0$.

Несколько более точный результат дает в нашем случае использование среднего абсолютного прироста за 1965—1970 гг., который равен 26,65 тыс. км $[(511,6 - 351,7) : 6]$. Интерполируемый уровень рассчитывается тогда по формуле

$$y_{1+t} = y_1 + \bar{P}t, \quad (9.37)$$

где 1 — номер начального (первого по порядку) уровня, использованного при расчете \bar{P} ;

\bar{P} — средний абсолютный прирост за весь рассматриваемый период;

t — длина периода между датой, на которую делается интерполяция $(1+t)$, и начальной датой, равная разности между номером интерполируемого уровня $(1+t)$ и номером начального уровня (1);

y_{1+t} — интерполируемый неизвестный уровень.

В нашем примере $y_1 = y_{1964} = 351,7$; $\bar{P}_{1965-1970} = 26,65$; $t = 1967-1964 = 3$. Отсюда

$$y_{1967} = y_{1964} + 3\bar{P} = 351,7 + 3 \times 26,65 = 431,7 \text{ (тыс. км)}.$$

Интерполировать по формуле (9.37) можно только в тех случаях, когда есть основания считать, что уровень ряда динамики в течение всего рассматриваемого периода изменяется в одном направлении, а цепные абсолютные приросты стабильны и мало отличаются один от другого.

Нередко, однако, уровень ряда динамики то повышается, то понижается или же растет, но очень неравномерно, так что цепные абсолютные приросты сильно отличаются один от другого, хотя и не имеют явной тенденции к росту или снижению, т. е. являются относительно стабильными. В таких случаях интерполяция целесообразно производить на основе аналитического выравнивания по прямой.

Рассмотрим следующий пример, где «неизвестные» уровни в действительности равны: $y_{1968} = 9,1$ млн. т, $y_{1969} = 9,6$ млн. т.

Таблица 9.24

ГОСУДАРСТВЕННЫЕ ЗАКУПКИ ОВОЩЕЙ В СССР

Годы	Закупки у	t	yt	t^2
1964	7,9	1	7,9	1
1965	7,7	2	15,4	4
1966	8,0	3	24,0	9
1967	9,5	4	38,0	16
1968	...	—	—	—
1969	...	—	—	—
1970	10,9	7	76,3	49
Итого	44,0	17	161,6	79

В 1965—1967 гг. закупки овощей увеличивались в среднем ежегодно на 0,53 млн. т, а в 1968—1970 гг. — на 0,47 млн. т $[(10,9 - 9,5) : 3]$. Следовательно, абсолютные приросты не имеют явной тенденции к увеличению или снижению, т. е. относительно стабильны. Однако они сильно колеблются от года к году ($P_{1965} = -0,2$; $P_{1966} = 0,3$; $P_{1967} = 1,5$). Поэтому произведем интерполяцию путем выравнивания ряда по прямой.

В отличие от обычного выравнивания здесь не заполняются графы « y_t » и « t^2 » по тем строкам, где отсутствуют уровни. Кроме того, отсчет времени (t) нужно вести по порядку, начиная с первого года (а не от центра ряда). При этом номера тех лет, уровни которых неизвестны, не записываются, но учитываются в нашем примере 1968 г. — пятый по счету, 1969 — шестой, следовательно, 1970 — седьмой).

Система нормальных уравнений в нашем примере имеет следующий вид:

$$5a_0 + 17a_1 = 44,0,$$

$$17a_0 + 79a_1 = 161,6.$$

Ее решение приводит к прямой $\hat{y} = 6,88 + 0,566t$.

Интерполированный по этой прямой уровень 1968 г. ($t = 5$) составляет 9,7 млн. т, а уровень 1969 г. ($t = 6$) — 10,3 млн. т.

В тех случаях, когда ценные абсолютные приросты более или менее равномерно возрастают, для интерполяции следует использовать выравнивание динамического ряда по параболе второго порядка.

Если есть основания считать, что ценные темпы роста мало отличаются один от другого, интерполяция производится на основе среднего темпа роста (\bar{T}_p):

$$y_{t+1} = y_t \times (\bar{T}_p)^t, \quad (9.38)$$

где обозначения (y_t, y_{t+1}, t) те же, что в формуле (9.37).

Так, численность специалистов с высшим и средним специальным образованием, занятых в народном хозяйстве СССР, составляла (на 15 ноября; тысяч человек): в 1965 г. — 12 066, в 1968 г. — 14 956. Среднегодовой темп роста равен 10,74%.

Предполагая, что за первый год этого трехлетия темп роста был равен этой средней величине, найдем уровень 1966 г. ($t = 1966 - 1965 = 1$):

$$y_{1966} = y_{1965} \times \bar{T}_p = 12 066 \times 1,074 = 12 959 \text{ (тыс. человек)}.$$

Это лишь на 35 тыс. человек (менее чем на 0,3%) отличается от действительной численности на 15/XI 1966 г. (12 924 тыс. человек).

Если ценные темпы роста относительно стабилизировались, т. е. не имеют тенденции к ускорению или замедлению, но сильно отличаются один от другого, интерполяцию лучше делать на основе аналитического выравнивания по показательной кривой ($\hat{y}_t = a_0 e^{ct}$).

В нашем примере уровень 1966 г., интерполированный по показательной кривой $\hat{y}_t = 11 320 \times 1,069^t$ (полученной с учетом уровня 1970 г. — 16 841 тыс. человек), составил 12 940 тыс. человек.

Приближенный расчет неизвестных уровней динамического ряда, лежащих за его пределами, т. е. в будущем (или в прошлом), называется экстраполяцией ряда динамики. При экстраполяции исходят из предположения, что характер динамики, т. е. та или иная закономерность развития, имевшая место в течение известного периода, сохранится также и в будущем (перспективная экстраполяция) или имела место и в прошлом (ретроспективная экстраполяция).

По сравнению с интерполяцией экстраполяция является менее обоснованной и надежной и дает еще более приближенные результаты. Это связано с тем, что при интерполяции рассчитываются уровни, лежащие внутри интервала времени, для которого выявлены определенная закономерность, а при экстраполяции — уровни, расположенные вне этого интервала, за его пределами, когда закономерность может оказаться уже иной, измениться.

Особое значение при экстраполяции имеют вопросы ее базы и сроков, т. е. о продолжительности периода, характерную для которого закономерность можно распространить на будущее, и о длине периода, на который можно распространить эту закономерность. В качестве базы для экстраполяции нельзя, очевидно, брать очень короткий период, так как он может оказаться недостаточно типичным из-за действия временных, переходящих условий и факторов. Нецелесообразно брать за базу и очень длинный период, так как условия развития явления с течением времени изменяются обычно все больше и больше. В результате то, что было характерно для динамики 15—20 лет назад, может оказаться совершенно не характерным в ближайшие годы. Оптимальное по длительности база при экстраполяции должна быть выбрана на основе теоретического анализа сущности явления с учетом конкретных исторических условий его развития. Что касается срока экстраполяции, то чем он короче, тем более точные результаты при прочих равных условиях даст экстраполяция, так как за короткий срок не

успеют сильно измениться условия, развития явления и характер его динамики.

Как и интерполяция, экстраполяция может производиться на основе среднего абсолютного прироста (формула 9.37), среднего темпа роста (формула 9.38), а также с помощью аналитического выравнивания по прямой или какой-либо кривой.

Так, выше было установлено, что в 1965—1970 гг. протяженность автомобильных дорог с твердым покрытием увеличивалась в среднем ежегодно на 26,65 тыс. км. Если эта тенденция сохранится, то к концу 1972 г. ($t = 1972 - 1964 = 8$) протяженность этих дорог должна была составлять около 565 тыс. км ($y_{1972} = y_{1964} + \bar{P}t = 351,7 + 26,65 \times 8$).

Если ряд динамики государственных закупок овощей (см. табл. 9.24) дополнить уровнями за 1968 г. (9,1 млн. т) и 1969 г. (9,6 млн. т) и выравнивать по прямой, ведя отсчет времени от центра ряда (для 1967 г. значение $t = 0$), получим следующее уравнение: $\hat{y}_t = 8,96 + 0,496t$. Подставив в него значение t для 1971 г., равное 4, найдем экстраполированный уровень закупок в 1971 г.: $y_{1971} = 10,94$ (млн. т). Фактически закупки овощей в этом году составили 11,5 млн. т.

В качестве исходных данных для экстраполяции на основе средних темпов роста используем численность специалистов с высшим и средним специальным образованием, занятых в народном хозяйстве СССР. На 15/XI 1965 г. она составляла 12 066 тыс. человек, а на 16/XI 1970 г. — 16 841 тыс. человек. Среднегодовой темп роста равен 10,69%. Если такой темп сохранился и в 1971 г., то в ноябре этого года численность специалистов должна была составить 18 003 тыс. человек ($16 841 \times 1,069$). По оценке ЦСУ СССР, на конец 1971 г. она составляла 17,9 млн. человек.

1. ОБЩЕЕ ПОНЯТИЕ ОБ ИНДЕКСАХ И ИНДЕКСНОМ МЕТОДЕ

Индексы в широком понимании Латинское слово *индекс* (*index*) в переводе означает «указатель», «показатель».

В экономике и статистике этот термин употребляется как в широком, так и в более узком, специальном понимании.

В широком смысле под индексом понимается относительный показатель, который характеризует соотношение уровня социально-экономического явления во времени, по сравнению с планом или в пространстве. Во всех этих случаях сопоставляются между собой уровни одноименных явлений, имеющих одинаковое экономическое содержание. Следовательно, индексами в широком смысле являются любые разновидности относительных величин трех видов: динамики (включая плановые темпы роста), степени выполнения плана и сравнения в пространстве. Так, например, если мы говорим, что производство электроэнергии в СССР в 1973 г. увеличилось по сравнению с 1940 г. почти в 19 раз, то эту относительную величину в широком смысле можно назвать индексом динамики производства электроэнергии. Еще пример. Допустим, план производства молока в каком-либо колхозе выполнен только на 98%, а план производства всей сельскохозяйственной продукции — на 105%. Эти относительные величины являются индексами выполнения плана производства продукции.

Индивидуальные и сводные индексы Изучаемые статистикой социально-экономические явления обычно имеют сложный характер, состоят из единиц разного вида.

Так, каждое промышленное предприятие выпускает, как правило, несколько видов продукции, различных (неоднородных) по своей натурально-вещественной форме. При этом некоторые виды продукции имеют свои разновидности и сорта (например, пальто зимние и демисезонные и т. п.). Различные виды продукции производят и сельскохозяйственные предприятия — колхозы и совхозы. Еще более широкий круг продукции производит все предприятия какой-либо отрасли промышленности или все колхозы и совхозы области, республики. Наконец, валовой общественный продукт страны в его материально-вещественной

форме представляет собой совокупность продуктов всех отраслей материального производства.

В соответствии с этим по кругу охвата единиц изучаемой совокупности индексы делятся на индивидуальные и сводные. Индивидуальный индекс характеризует соотношение уровней явления по отдельному виду единиц совокупности (например, соотношение объема производства одного вида продукции или себестоимости единицы данного вида продукции). Так, например, приведенные выше индексы динамики производства электроэнергии и индекс выполнения плана производства молока являются индивидуальными.

Сводный индекс характеризует соотношение уровней сложного явления, состоящего из нескольких различных видов единиц (например, объема производства нескольких видов продукции, имеющих различную натурально-вещественную форму). Так, сводным индексом является приведенный выше индекс выполнения плана производства всей сельскохозяйственной продукции колхоза.

Если изучаемая совокупность состоит из нескольких групп единиц, то сводные индексы, каждый из которых характеризует соотношение уровней отдельной группы единиц, являются групповыми (субиндексами), а сводный индекс, охватывающий все группы, является общим индексом. Так, индексы, характеризующие изменение объема продукции растениеводства и объема продукции животноводства, являются сводными групповыми индексами, а индекс, характеризующий изменение объема всей сельскохозяйственной продукции колхоза — общий индексом. Если группы единиц делятся далее на подгруппы, то кроме субиндексов I порядка могут быть исчислены субиндексы II и других порядков (например, наряду с субиндексами I порядка, охватывающими продукты растениеводства и животноводства, субиндексы II порядка, охватывающие продукты отдельных отраслей растениеводства и животноводства (полеводства, садоводства и т. п.).

Индексы объемных и качественных показателей Показатель, соотношение уровней которого характеризует индекс, называется индексирваемым показателем. Индексирваемые показатели могут быть двоякого рода. Одни из них характеризуют общий, суммарный размер (объем) того или иного явления, например количество (физический объем) продукции в натуральном выражении, число рабочих, общие затраты времени (труда) на производство продукции, общая себестоимость продукции и т. п. Такие показатели называются объемными. Они представляют собой или численность тех или иных единиц или общий объем какого-либо признака. Объемные показатели получают как итог непосредственного подсчета или суммирования и являются основными (исходными, первичными).

Другие показатели характеризуют уровень явления в расчете на ту или иную единицу совокупности: выработка продукции за

единицу времени (или на одного работника), затраты рабочего времени (труда) на единицу продукции, себестоимость единицы продукции и др. Такие показатели называются **качественными**. Они исчисляются путем деления объемных показателей, т. е. носят расчетный, вторичный характер. Качественные показатели измеряют не общий объем, а **интенсивность, эффективность** явления или процесса. Как правило, они являются либо **средними** либо **относительными** величинами.

Индексная символика

При использовании индексного метода применяется определенная символика, т. е. система условных обозначений, позволяющих строить и записывать формулы индексов в сокращенном виде.

Каждый индексируемый показатель, соотношение уровней которого характеризует индекс, принято обозначать определенной буквой (обычно латинской). В дальнейшем будет использоваться также условные обозначения:

Q — количество (объем) произведенной продукции и количество проданных товаров данного вида в натуральном выражении (объемный показатель);

p — цена единицы продукции или товара (качественный показатель);

z — себестоимость единицы продукции (качественный показатель);

T — общие затраты времени (труда) на производство продукции данного вида, измеряемые в человеко-часах, человеко-днях и т. п. В некоторых случаях этой же буквой (T) будет обозначаться **среднеспособная численность работников (или рабочих)** за тот или иной период, так как она численно равна количеству отработанных человеко-месяцев, человеко-лет и т. п. (объемный показатель).

На основе этих показателей могут быть получены другие:

pQ — **общая стоимость**¹ произведенной продукции данного вида или **общая стоимость проданных товаров** данного вида (товарооборот, **кассовая выручка**) — объемный показатель;

zQ — **общая себестоимость продукции** данного вида, т. е. **денежные затраты на ее производство** (объемный показатель);

$t = \frac{T}{Q}$ — **затраты рабочего времени (труда) на производство единицы продукции** данного вида, т. е. **трудоемкость единицы продукции** (качественный показатель);

$q = \frac{Q}{T}$ — **производство продукции** данного вида за единицу рабочего времени (за 1 человеко-час, человеко-день, человеко-месяц

и т. д.) или **производство продукции на одного среднеспособного работника или рабочего**, т. е. **уровень производительности труда** (качественный показатель).

Показатели за период, с которым производится сравнение (**базисный период**), сопровождаются подстрочной цифрой «0», показатели за **сравниваемый (отчетный, текущий) период** — цифрой «1», а если сравниваются данные за несколько периодов, то указываются либо их порядковые номера, либо конкретные хронологические периоды (годы, кварталы и т. п.). Плановые показатели снабжаются знаком «пл».

Индивидуальные индексы обозначаются буквой i и снабжаются подстрочным значком индексируемого показателя. Так, i_Q означает индивидуальный индекс объема продукции отдельного вида или количества (объема) проданного товара определенного вида; i_z — индивидуальный индекс себестоимости единицы продукции данного вида; i_{pQ} — индивидуальный индекс общей себестоимости продукции отдельного вида, т. е. затрат на ее производство; i_t — индивидуальный индекс производительности труда и т. п.

Сводный индекс обозначается буквой I и также сопровождается подстрочным значком индексируемого показателя. Например, I_p — сводный индекс цен; I_{pQ} — сводный индекс общей стоимости продукции или общей стоимости проданных товаров, т. е. товарооборота; I_{zQ} — сводный индекс общей себестоимости продукции (затрат на ее производство); I_t — сводный индекс трудоемкости единицы продукции нескольких видов; I_T — сводный индекс общих затрат времени (труда) или же сводный индекс численности рабочих (или работников) и т. д.

Индивидуальный индекс, характеризующий соотношение уровней явления по отдельному виду единиц совокупности, является обычной относительной величиной: либо относительной величиной динамики, т. е. фактическим или плановым темпом роста, либо степенью выполнения плана, либо, наконец, относительной величиной сравнения в пространстве (территориальный индекс).

Индивидуальные индексы не являются собственно индексами, т. е. они могут быть названы индексами только при широком понимании этого термина (в целях единства методики и терминологии).

В соответствии с приведенными выше условными обозначениями различные варианты формулы индивидуального индекса объема продукции в натуральном выражении, т. е. физического объема продукции, будут таковы:

$$i_Q = \frac{Q_1}{Q_0} \text{ — индекс динамики (темп роста);} \quad (10.1)$$

$$i_Q = \frac{Q_{пл}}{Q_0} \text{ — индекс планового задания (плановый темп роста);} \quad (10.2)$$

¹ Здесь и далее термин «стоимость» используется не как категория политической экономии, а как величина, являющаяся произведением цены товара или продукции на их количество.

$$i_q = \frac{Q_1}{Q_{пл}} — \text{индекс (степень) выполнения плана;} \quad (10.3)$$

$$i_q = \frac{Q_A}{Q_B} — \text{территориальный индекс, т. е. соотношение объема продукции на предприятии или в районе (области, республике и т. п.) А и объема продукции на предприятии (районе и т. д.) В.} \quad (10.4)$$

Аналогично строятся индивидуальные индексы и других показателей. Например, индивидуальный индекс динамики себестоимости единицы продукции имеет такой вид:

$$i_z = \frac{z_1}{z_0}, \quad (10.5)$$

а индивидуальный индекс выполнения плана по себестоимости единицы продукции — такой:

$$i_z = \frac{z_1}{z_{пл}}. \quad (10.6)$$

Индивидуальный индекс планового задания по трудоемкости единицы продукции строится так:

$$i_t = \frac{t_{пл}}{t_0}, \quad (10.7)$$

а индивидуальный индекс динамики производительности труда — так:

$$i_q = \frac{q_1}{q_0}. \quad (10.8)$$

Уровень производительности труда, т. е. производство продукции за единицу рабочего времени ($q = Q : T$), и трудоемкость единицы продукции, т. е. затраты времени на единицу продукции ($t = T : Q$), показатели обратные, так что $tq = 1$, $t = \frac{1}{q}$, $q = \frac{1}{t}$. Поэтому индивидуальный индекс производительности труда также представляет собой величину, обратную индивидуальному индексу трудоемкости единицы продукции:

$$i_q = q_1 : q_0 = \frac{1}{t_1} : \frac{1}{t_0} = \frac{t_0}{t_1} = \frac{1}{i_t}, \quad (10.9)$$

где $i_t = t_1 : t_0$ — индивидуальный индекс трудоемкости единицы продукции.

Поскольку индивидуальные индексы являются обычными относительными величинами, для них сохраняется силу взаимосвязь между плановым темпом роста, степенью выполнения плана и фактическим темпом роста:

$$\frac{z_{пл}}{z_0} \cdot \frac{z_1}{z_{пл}} = \frac{z_1}{z_0}, \text{ или } i_{\text{план. задания}} \cdot i_{\text{выпол. плана}} = i_{\text{динамики}}.$$

Для индивидуальных индексов динамики и планового задания сохраняется силу и взаимосвязь цепных и базисных темпов роста. Так,

$$\frac{q_1}{q_0} \cdot \frac{q_2}{q_1} \cdot \frac{q_3}{q_2} = \frac{q_3}{q_0},$$

где в левой части равенства — произведение трех цепных индексов, а в правой — базисный индекс, характеризующий темп роста производительности труда в третьем периоде по сравнению с базисным.

Наконец, для индивидуальных индексов всех четырех типов (динамики, планового задания и т. д.) справедлива и следующая взаимосвязь: если произведение двух или нескольких показателей представляет собой новый показатель, имеющий реальный экономический смысл, то индекс произведения нескольких показателей равен произведению индексов этих показателей-составителей. Так, применительно к общим затратам времени (T), трудоемкости единицы продукции (t) и объему продукции (Q) имеем:

$$T = tQ; \quad \frac{T_1}{T_0} = \frac{t_1 Q_1}{t_0 Q_0} = \frac{t_1}{t_0} \cdot \frac{Q_1}{Q_0} \text{ или } i_T = i_t \cdot i_Q. \quad (10.10)$$

Аналогично

$$Q = qT; \quad \frac{Q_1}{Q_0} = \frac{q_1 T_1}{q_0 T_0} = \frac{q_1}{q_0} \cdot \frac{T_1}{T_0} \text{ или } i_Q = i_q \cdot i_T. \quad (10.11)$$

Такого рода взаимосвязи позволяют производить так называемое косвенное индексирование, т. е. по двум или нескольким известным индексам находить связанный с ним индекс, числовое значение которого неизвестно. Если, например, известно, что объем произведенной продукции возрос на 20%, а общие затраты времени на ее производство увеличились на 8%, то $i_Q = 1,2$, $i_T = 1,08$ и согласно (10.10) $1,08 = i_t \times 1,2$, откуда $i_t = 1,08 : 1,2 = 0,9$, или 90%, т. е. трудоемкость единицы продукции снизилась на 10%. В то же время согласно формуле (10.11) $1,2 = i_q \times 1,08$, откуда $i_q = 1,2 : 1,08 = 1,111$, или 111,1%, т. е. производительность труда повысилась на 11,1%.

Экономические явления и характеризующие их показатели могут быть соизмеримыми, т. е. иметь общую меру, и несоизмеримыми. Так, объемы (количества) одного и того же вида продукции, произведенные на разных предприятиях, соизмеримы, объемы же разных видов продукции или разных товаров несоизмеримы.

Соизмеримые объемные показатели можно непосредственно суммировать, несоизмеримые — нельзя. Например, нельзя складывать килограммы хлеба с литрами молока, метрами ткани и парами обуви. Несоизмеримость и невозможность непосредственного суммирования объясняется здесь не столько различием натуральных единиц измерения, сколько различием потребительных свойств

отдельных товаров, их неодинаковой натурально-вещественной формой. Количество проданного мяса измеряется в тех же весовых единицах, что и количество рыбы, однако непосредственное суммирование килограммов мяса и рыбы все же экономически бессмысленно, так как полученная величина представляла бы собой в прямом смысле «ни рыбу, ни мясо». Причиной несоизмеримости является здесь неоднородность, различие натурально-вещественной формы.

В отличие от этого некоторые другие объемные показатели всегда соизмеримы и могут поэтому непосредственно суммироваться. К ним относятся, например, общие затраты времени на производство различных видов продукции (T), общая себестоимость различных видов продукции (ΣQ), общая стоимость проданных товаров (кассовая выручка, товарооборот — PQ) и др. Поэтому построение и исчисление сводных индексов таких соизмеримых объемных показателей не представляют никаких трудностей. Эти индексы, как и индивидуальные, представляют собой обычные относительные величины и могут быть названы индексами только в широком их смысле:

$$I_T = \frac{\Sigma T_1}{\Sigma T_0}, \quad (10.12) \quad I_{PQ} = \frac{\Sigma (PQ)_1}{\Sigma (PQ)_0} = \frac{\Sigma p_1 Q_1}{\Sigma p_0 Q_0}. \quad (10.13)$$

Иначе обстоит дело с построением сводных индексов несоизмеримых объемных показателей (объемы разных видов продукции или товаров). В результате невозможности непосредственного суммирования в этом случае для получения общего итога нужно предварительно привести отдельные виды продукции или товаров к соизмеримому виду, т. е. соизмерить их с помощью тех или иных коэффициентов соизмерения, обеспечив тем самым возможность последующего суммирования и построения и расчета сводного индекса.

Таким образом, построение сводных индексов несоизмеримых объемных показателей требует использования специальных приемов, которые составляют специфику индексного метода. В связи с этим сводные индексы несоизмеримых объемных показателей являются не обычными, а особыми относительными величинами, они имеют более сложную методику построения и расчета. В отличие от обычных относительных величин они являются индексами в собственном смысле слова, индексами в узком понимании, так как их построение и исчисление связано с использованием специфических приемов, составляющих суть индексного метода (эти приемы будут рассмотрены ниже).

К собственно индексам (индексам в узком понимании) относятся также все сводные индексы качественных показателей. Допустим, нас интересует изменение общего уровня цен на проданные товары. Хотя цены разных товаров соизмеримы, однако их непосредственное суммирование, т. е. суммирование без учета количества каждого товара, даст величину, лишенную самостоятельного практического значения. Действительно, сложив, напри-

мер, цены 1 кг хлеба, 1 пачки чая, 1 м ткани, 1 пары обуви и т. д., мы получим общую стоимость совершенно произвольного набора товаров. Между тем ясно, что нас интересует изменение общего уровня цен не того или иного произвольного набора товаров, а уровня цен фактически проданного их количества, образующего вполне определенную, реально существующую совокупность.

Простое суммирование цен за ту или иную единицу каждого товара дает такую суммарную величину, отдельные слагаемые которой несравнимы, так как оторваны от конкретной совокупности единиц, к которым они относятся. Следовательно, при построении и исчислении сводных индексов качественных показателей возникает необходимость обеспечения сравнимости с помощью специальных приемов, специфичных для индексного метода. Поэтому сводные индексы качественных показателей также являются собственно индексами.

Проблема обеспечения сравнимости (сопоставимости) отдельных элементов сложного явления впервые в истории статистики была решена при построении сводного индекса цен, характеризующего динамику общего уровня цен разных товаров в виде одной относительной величины. Было предложено суммировать не просто цены тех или иных единиц товаров (как это делалось ранее — Σp), а общие стоимости проданных (или потребленных) товаров, т. е. произведения цены данного товара на его количество (ΣpQ). Таким образом, первая специфическая особенность индексного метода и основанных на нем собственно индексов состоит в том, что изменение изучаемого явления рассматривается не изолированно, а во взаимосвязи с другим явлением: цены — во взаимосвязи с количеством товаров, себестоимость и трудоемкость единицы различных видов продукции — во взаимосвязи с количеством этих видов продукции и т. п.

Умножение индексированного качественного показателя на непосредственно связанный с ним объемный показатель в известной степени аналогично умножению вариантов признака на их веса при расчете средней арифметической взвешенной. В обоих случаях умножением заменяется сложение одинаковых слагаемых — значений признака по группе единиц совокупности. Умножая цену товара на его количество, мы придаем этому товару соответствующий вес в общей стоимости всех товаров. Следовательно, показатели-сомножители, связанные с индексированным показателем, при построении собственно индексов играют роль статистических весов.

Однако умножение индексированного показателя на связанные с ним значения другого показателя (веса) еще не решает проблему построения сводного индекса. Допустим, что имеются следующие данные по колхозному рынку (для простоты взяты только два товара; см. табл. 10.1).

Как видно из приведенных данных, цены обоих товаров в текущем периоде по сравнению с базисным снизились, и нас интересует, как изменился уровень цен в целом по обоим товарам.

Таблица 10.1

КОЛИЧЕСТВО И ЦЕНЫ ПРОДАННЫХ ТОВАРОВ

Товары	Исходные данные				Расчет		
	базисный период		текущий период		стоимость проданных товаров (руб.)		стоимость товаров в отчетном периоде по ценам базисного периода (руб.)
	продано (кг)	цена 1 кг (руб.)	продано (кг)	цена 1 кг (руб.)	базисный период	текущий период	
	Q_0	P_0	Q_1	P_1	$P_0 Q_0$	$P_1 Q_1$	$P_0 Q_1$
Мясо	230	2,80	300	2,5	644	750	840
Картофель	624	0,25	640	0,2	156	128	160
Итого	—	—	—	—	800	878	1 000

Умножив цену каждого товара на его количество, проданное по этой цене, мы получим общую стоимость всего проданного товара за каждый период (товарооборот — $P_0 Q_0$ и $P_1 Q_1$). В каждом периоде стоимость мяса сравнима со стоимостью картофеля, так как обе они представляют всю фактическую выручку за проданный товар. Стоимость обоих проданных товаров, т. е. сумма произведений цен на количество ($\Sigma P_0 Q_0$ и $\Sigma P_1 Q_1$) имеет здесь ясный и реальный экономический смысл.

Однако если мы хотим определить, как изменился общий уровень цен, то, очевидно, при этом нельзя сопоставлять общую стоимость товаров, проданных в текущем периоде ($\Sigma P_1 Q_1$), со стоимостью товаров, проданных в базисном периоде ($\Sigma P_0 Q_0$), так как эти суммы произведений отражают не только изменение интересующего нас *индексируемого* показателя — цен, но и изменение *весов* — количества проданных товаров.

Следовательно, с точки зрения поставленной задачи — дать характеристику только изменения уровня цен — данные суммы произведений несравнимы между собой. Чтобы сделать стоимости товаров двух периодов сравнимыми, нужно сначала по ценам одного, а затем по ценам другого периода оценить одну и ту же совокупность товаров, один и тот же их набор — количество товаров, проданных либо в базисном, либо в текущем периоде. Только в этом случае сравнение полученных сумм произведений даст индекс, отражающий изменение только *индексируемого* показателя — цен. Изменение же весов (Q) будет в этом индексе устранено, элиминировано, так как в числителе и в знаменателе индекса веса будут зафиксированы на уровне одного и того же периода. Так, если в сводном индексе цен зафиксировать веса в числителе и знаменателе на уровне базисного периода, получим:

$$I_p = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \quad (10.14)$$

если же веса зафиксировать на уровне текущего периода, индекс цен будет иметь такой вид:

$$I_p = \frac{\sum P_1 Q_1}{\sum P_0 Q_1} \quad (10.15)$$

Произведя расчет сумм произведений (см. табл. 10.1) и подставив их в формулу (10.15), получим:

$$I_p = \frac{878}{1000} = 0,878, \text{ или } 87,8\%,$$

т. е. в текущем периоде по сравнению с базисным цены снизились на 12,2%.

Устранение влияния изменения весов путем их фиксирования в числителе и знаменателе индекса на одном и том же уровне — вторая специфическая особенность собственно индексов и индексного метода.

Специфические особенности индексного метода — рассмотрение индексируемого показателя во взаимосвязи с другим (статистическое взвешивание) и элиминирование изменения весов — используются и при построении сводных индексов несоизмеримых объемных показателей. Эти показатели рассматриваются во взаимосвязи с теми или иными качественными показателями, которые играют роль весов и фиксируются в числителе и знаменателе индекса на уровне одного и того же периода. Например, при построении сводного индекса объема проданных товаров в качестве весов (одновременно играющих роль коэффициентов соизмерения разных товаров) можно использовать цены, фиксирование которых на уровне базисного периода даст следующую формулу:

$$I_Q = \frac{\sum Q_1 P_0}{\sum Q_0 P_0} \quad (10.16)$$

В нашем примере (табл. 10.1) $I_Q = 1000 : 800 = 1,25$, или 125%, т. е. объем проданных товаров в сравнимых ценах базисного периода возрос на 25%.

Задачи, решаемые с помощью индексов

Рассматривая собственно индексы и возникающие при их построении проблемы, мы ставили перед этими индексами задачу дать сравнительную характеристику уровней сложного явления, состоящего из единиц разного вида. Так, перед сводным индексом цен ставилась задача показать, как изменился общий уровень цен, т. е. измерить динамику цен нескольких различных товаров в виде одного обобщающего показателя. Исторически собственно индексы появились как результат решения именно этой экономической задачи — задачи обобщения, синтеза динамики отдельных элементов сложного явления в одном обобщающем показателе — сводном индексе.

Однако впоследствии собственно индексы стали использоваться и для решения другой экономической задачи — для анализа

влияния изменения отдельных показателей-факторов на изменение экономического показателя, представляющего функцию этих факторов-аргументов. Так, если общая стоимость проданных товаров (товарооборот — ΣpQ) есть функция их цен (p) и количества (объемов — Q), то ставится задача оценить, измерить роль каждого из этих двух факторов в изменении товарооборота: определить, как он изменился отдельно за счет изменения каждого фактора. При этом индексы, используемые для решения подобных аналитических задач, также строятся с использованием специфических особенностей индексного метода — взвешивания и элиминирования изменения весов, хотя другие методологические проблемы построения индексов как показателей анализа в ряде случаев могут и должны решаться иначе, чем при построении индексов только в качестве обобщающих, синтетических показателей.

Некоторые собственно индексы, как будет показано далее, выполняют лишь функцию синтеза, обобщения, другие — только функции анализа; некоторые же индексы при определенных условиях могут одновременно решать обе задачи. В связи с этим может быть дано такое определение собственно индексов, охватывающее все их виды: **Собственно индекс** — это относительный показатель особого рода, в котором уровни социально-экономического явления рассматриваются во взаимосвязи с другим (или другими) явлением, изменение которого при этом элиминировано.

Показатели, связанные с индексированным показателем, используются в качестве весов индекса, а взвешивание и элиминирование изменения весов (их фиксирование в числителе и знаменателе индексного отношения на одном и том же уровне) составляют специфику собственно индексов и индексного метода.

Основы советской теории индексов были заложены в конце 20-х — начале 30-х годов. Большие заслуги в этой области принадлежат академику С. Г. Струмишину, члену-корреспонденту АН СССР В. Н. Старовскому и др. Разработка теории индексов (методология построения территориальных индексов, анализа влияния отдельных факторов на сложные экономические явления и т. п.) успешно ведется и в настоящее время (Перегрудов Н. В. Теоретические вопросы индексного анализа, 1960; Казинец Л. С. Теория индексов (основные вопросы), 1963 г. и др.). Однако некоторые вопросы остаются дискуссионными и до настоящего времени.

2. СВОДНЫЕ ИНДЕКСЫ СРАВНЕНИЯ УРОВНЕЙ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ

Построение индексов качественных показателей в агрегатной форме

Каждый качественный показатель связан с тем или иным объемным показателем, в расчете на единицу измерения которого он исчисляется (или к единице измерения которого относится). Так, цена единицы товара (p) связана с проданным его количеством, объемом (Q); с объемом произведенной продукции связаны такие ка-

чественные показатели, как цена (p), себестоимость (z) и трудоемкость ($t = T : Q$) единицы продукции; именно поэтому непосредственное суммирование значений качественного показателя у различных товаров или продуктов, т. е. суммирование без учета их количества ($\Sigma p, \Sigma z$ и т. д.), дает величину, лишенную самостоятельного экономического значения и смысла, а сводные индексы качественных показателей не могут строиться как отношение таких сумм ($\Sigma p_1 : \Sigma p_0$ и т. п.).

Сводные индексы качественных показателей должны характеризовать не их изменение вообще, а изменение цен, себестоимости или трудоемкости вполне определенного количества товаров или продукции, т. е. *определенного их набора*. Это достигается тем, что, во-первых, отдельные уровни индексированного качественного показателя умножаются на значения связанного с ним объемного показателя, играющие роль весов, и, во-вторых, эти веса в числителе и знаменателе индекса фиксируются на одном и том же уровне. Сопоставление сумм таких произведений, в которых индексированный качественный показатель взят на разных уровнях, а веса — на одинаковом уровне, и дает сводный индекс качественного показателя в агрегатной форме¹. Так, агрегатный индекс динамики цен, как было показано выше, может быть построен с весами, зафиксированными на уровне либо текущего ($I_p = \frac{\Sigma p_1 Q_1}{\Sigma p_0 Q_1}$), либо базисного периода ($I_p = \frac{\Sigma p_1 Q_0}{\Sigma p_0 Q_0}$). Аналогично может быть построен и агрегатный индекс динамики себестоимости единицы продукции:

$$I_z = \frac{\Sigma z_1 Q_1}{\Sigma z_0 Q_1}; \quad (10.17)$$

$$I_z = \frac{\Sigma z_1 Q_0}{\Sigma z_0 Q_0}. \quad (10.18)$$

Агрегатный индекс выполнения плана по себестоимости единицы продукции может быть построен либо с весами текущего (отчетного) периода, либо с плановыми весами:

$$I_z = \frac{\Sigma z_1 Q_1}{\Sigma z_{пл} Q_1}; \quad (10.19)$$

$$I_z = \frac{\Sigma z_1 Q_{пл}}{\Sigma z_{пл} Q_{пл}}. \quad (10.20)$$

Наконец, агрегатный индекс планового задания по себестоимости может быть взвешен по плановым и по базисным весам:

$$I_z = \frac{\Sigma z_{пл} Q_{пл}}{\Sigma z_0 Q_{пл}}; \quad (10.21)$$

$$I_z = \frac{\Sigma z_{пл} Q_0}{\Sigma z_0 Q_0}. \quad (10.22)$$

¹ От слова «агрегат» — соединение отдельных частей в одно целое.

Если в формулах (10.17) — (10.22) в качестве индексируемой величины вместо себестоимости (z) принять трудоемкость единицы продукции (t), то получим соответствующие агрегатные индексы трудоемкости. Так, агрегатный индекс динамики трудоемкости единицы продукции при фиксировании весов на уровне текущего периода имеет такой вид:

$$I_t = \frac{\sum t_1 Q_1}{\sum t_0 Q_1} \quad (10.23)$$

Так как $tQ = T$, то числитель этого индекса представляет собой общие затраты времени в текущем периоде. Знаменатель — условная величина, которая показывает, сколько было бы затрачено времени на производство продукции текущего периода, если бы затраты времени на единицу продукции сохранялись на базисном уровне (t_0).

Основной задачей при построении сводных индексов качественных показателей является экономически обоснованный выбор того конкретного набора товаров или продукции, который следует использовать в качестве весов, чтобы с помощью этого индекса можно было решить конкретную экономическую задачу. Иными словами, главной задачей является *выбор уровня, на котором нужно зафиксировать веса индекса*.

Основной задачей перед сводным индексом динамики качественного показателя ставится задача охарактеризовать не только относительное изменение уровня, но и абсолютную величину того *экономического эффекта*, который получен в текущем периоде в результате этого изменения. Так, индексы цен и себестоимости единицы продукции должны не только показать, на сколько процентов снизился (или повысился) уровень цен или себестоимости, но и охарактеризовать полученный за счет этого экономический эффект — сумму экономии покупателей за счет снижения цен (или сумму их дополнительных расходов, если цены повысились), сумму экономии, полученной предприятием в результате снижения себестоимости, и т. п.

Такая постановка задачи приводит к индексам динамики *качественных* показателей с весами *текущего* периода: во-первых, нас интересует изменение себестоимости или трудоемкости той продукции, которая выпущена в настоящее время, а не в прошлом, во-вторых, экономический эффект должен быть увязан с фактическими результатами текущего, а не какого-либо предыдущего периода.

Например, имеются следующие данные (см. табл. 10.2).

Себестоимость единицы продукции каждого вида снизилась. Следовательно, в текущем периоде за счет этого получена экономия. Так, при производстве каждого стула фабрика сэкономила 0,2 руб. ($z_1 - z_0 = 2,8 - 3,0$). Всего в текущем периоде было произведено 12 тыс. стульев, так что общая сумма экономии за счет снижения себестоимости стульев составила 2400 руб. [$(z_1 - z_0) Q_1 = z_1 Q_1 - z_0 Q_1 = 33\,600 - 36\,000$]. Аналогично могут быть исключены сум-

Виды продукции	Базисный период			Текущий период			Затраты на объем продукции текущего периода при базисной себестоимости единицы продукции
	произведено (штук)	затраты на производство (руб.)	себестоимость единицы (руб.)	произведено (штук)	затраты на производство (руб.)	себестоимость единицы (руб.)	
	Q_0	$z_0 Q_0$	z_0	Q_1	$z_1 Q_1$	z_1	$z_0 Q_1$
Стулья	9 000	27 000	3	12 000	33 600	2,8	36 000
Стол	1 800	36 000	20	2 000	38 000	19,0	40 000
Диваны	460	23 000	50	480	23 280	48,5	24 000
Итого	—	86 000	—	—	94 880	—	100 000

мы, сэкономленные при выпуске других видов продукции. В целом по всем видам продукции сумма экономии составляет 5120 руб. ($\sum z_1 Q_1 - \sum z_0 Q_1 = 94\,880 - 100\,000$).

Таким образом, *реальная, фактически достигнутая в текущем периоде экономия за счет снижения себестоимости единицы продукции выражается как разность тех самых показателей, путем деления которых исчисляется индекс себестоимости, взвешенный по весам текущего периода*.

В нашем примере $I_z = \frac{\sum z_1 Q_1}{\sum z_0 Q_1} = \frac{94\,880}{100\,000} = 0,949$, или 94,9%, т. е. себестоимость снизилась на 5,1%.

В связи с этим в индексе динамики себестоимости веса обычно и фиксируются именно на уровне текущего периода. Это относится и ко всем другим агрегатным индексам динамики качественных показателей. Взвешивание по весам текущего периода увязывает эти индексы с абсолютной величиной реального экономического эффекта, достигнутого в результате изменения уровня качественного показателя.

Если индекс динамики себестоимости взвесить по весам базисного периода, то индекс охарактеризует ее изменение применительно к соотношению объемов выпуска отдельных видов продукции в базисном периоде, а разность между числителем и знаменателем покажет, каков был бы экономический эффект от снижения себестоимости, если бы выпуск продукции в текущем периоде оставался таким же, как в базисном. Иными словами, эта разность будет выражать не фактически полученный, а условный эффект. В нашем примере

$$I_z = \frac{\sum z_1 Q_0}{\sum z_0 Q_0} = \frac{2,8 \times 9\,000 + 19 \times 1\,800 + 48,5 \times 460}{86\,000} = \frac{81\,710}{86\,000} = 0,950.$$

Этот результат несколько отличается от предыдущего. Следовательно, *фиксирование весов устраняет лишь влияние их изме-*

нения, но не освобождает индекс от влияния весов вообще. Это обстоятельство и требует экономически обоснованного выбора весов индекса.

При построении индексов планового задания по себестоимости и трудоемкости единицы продукции экономические соображения приводят к взвешиванию этих индексов по весам плана (формула 10.21). Что же касается индексов выполнения плана, то их исчисление целесообразно производить в двух вариантах: 1) в расчете на продукцию, произведенную в отчетном (текущем) периоде, и 2) применительно к продукции, предусмотренной планом. Первый вариант индекса покажет степень выполнения плана применительно к фактически выпущенной в отчетном периоде продукции. Отношение же этого индекса к индексу, взвешенному по весам плана, будет показывать, как повлияло на степень выполнения плана отклонение фактического ассортимента¹ от того, который предусматривался в плане. Отношение двух взвешенных индексов называется индексом влияния ассортиментных сдвигов. Применительно к выполнению плана по себестоимости индекс ассортиментных сдвигов имеет такой вид:

$$I_{\text{ассорт. сдвигов}} = \frac{\sum z_1 Q_1}{\sum z_{\text{пл}} Q_1} : \frac{\sum z_1 Q_{\text{пл}}}{\sum z_{\text{пл}} Q_{\text{пл}}} \quad (10.24)$$

В приведенном выше примере (табл. 10.2) все три вида продукции выпускались как в базисном, так и в текущем периоде. Однако нередко отдельные продукты, которые выпускали в базисном периоде, не выпускают в текущем или, наоборот, продукты, которые не производили в базисном периоде, начинают производить в текущем периоде. Так, в нашем примере выпуск обычных диванов, производившихся в базисном периоде, в отчетном периоде мог быть прекращен и заменен выпуском нового вида продукции — диванов-кроватей.

В таких случаях возникает вопрос о том, как исчислять сводные индексы качественных показателей: с охватом всех видов продукции (или товаров) или как-то иначе? Поскольку индекс качественного показателя должен характеризовать изменение его уровня применительно ко вполне определенному кругу единиц, то очевидно, что этот круг единиц должен быть одинаковым для обоих сопоставляемых периодов. Такой круг единиц (набор видов продукции или товаров) называется сравнимым или сопоставимым кругом. В него входят только те продукты или товары, которые производились или продавались как в базисном, так и в текущем периоде, и не входят те, которые производились (продавались) только в базисном или только в текущем периоде.

¹ Под ассортиментом здесь понимается определенное соотношение между выпуском различных видов продукции.

Построение агрегатного индекса производительности труда

В некоторых случаях умножение индексируемого качественного показателя на связанный с ним объемный показатель, т. е. взвешивание, дает такой объемный показатель, элементы которого являются несоизмеримыми и не могут суммироваться. Это усложняет построение сводного индекса в агрегатной форме. Рассмотрим такой случай на примере индекса производительности труда.

Уровень производительности труда измеряется выработкой продукции за единицу времени (или выработкой продукции в расчете на одного работника): $q = Q : T$. Чтобы построить сводный индекс производительности труда в агрегатной форме, в качестве весов нужно принять общие затраты времени (или число работников) — T . Однако в результате такого взвешивания мы получим объем продукции в натуральном выражении: $qT = Q$. Если продукция разнородная, то ее различные виды несоизмеримы и не могут непосредственно суммироваться, поэтому построение сводного агрегатного индекса обычным путем ($I_q = \frac{\sum q_1 T_1}{\sum q_0 T_0}$) неприемлемо.

В связи с этим сводный агрегатный индекс производительности труда строится иначе: как обратная величина сводного агрегатного индекса трудоемкости единицы продукции. При этом исходят из того, что если уровень производительности труда (q) и трудоемкость единицы продукции (t) являются обратными показателями, то и их индексы, в том числе сводные, тоже являются величинами обратными:

$$I_q = \frac{1}{I_t} = \frac{\sum t_0 Q_1}{\sum t_1 Q_1} \quad (10.25)$$

Например, имеются следующие данные.

Таблица 10.3

ПРОИЗВОДСТВО ПРОДУКЦИИ И ЗАТРАТЫ ТРУДА В КОЛХОЗЕ

Виды продукции	Исходные данные				Расчет	
	1970 г.		1973 г.		затратено человеко-дней на 1 ц в 1970 г.	было бы затратено человеко-дней на продукцию в 1973 г. при трудоемкости 1970 г.
	произведено (тыс. ц)	затратено человеко-дней (тыс.)	произведено (тыс. ц)	затратено человеко-дней (тыс.)		
	Q_0	T_0	Q_1	T_1	$t_0 = T_0 : Q_0$	$t_0 Q_1$
Зерно	30	7,5	36	7,6	0,25	9,0
Молоко	10	17,0	11	17,4	1,70	18,7
Итого	—	24,5	—	25,0	—	27,7

Для исчисления сводного индекса производительности труда по формуле (10.25) найдем сначала затраты времени на 1 ц про-

дукции в 1970 г., т. е. базисную трудоемкость (t_0). Умножив ее на объем продукции 1973 г., получим условные затраты времени t_0Q_1 (последняя графа таблицы). Учитывая, что $t_1Q_1 = T_1$, исчислим сводный индекс:

$$I_q = \frac{\sum t_0 Q_1}{\sum t_1 Q_1} = \frac{27,7}{25,0} = 1,108, \text{ или } 110,8\%.$$

Таким образом, в целом по обоим видам продукции производительность труда повысилась на 10,8%. В результате этого вместо 27,7 тыс. человеко-дней, которые потребовались бы для производства продукции в 1973 г., если бы трудоемкость 1 ц оставалась на уровне 1970 г., фактически в 1973 г. было затрачено только 25 тыс. человеко-дней. Следовательно, за счет снижения трудоемкости, т. е. за счет роста производительности труда, было сэкономлено 2,7 тыс. человеко-дней, в том числе при производстве зерна — 1,4 тыс. ($T_1 - t_0 Q_1 = 7,6 - 9,0$) и при производстве молока — 1,3 тыс. (17,4 — 18,7).

Построение агрегатных индексов объемных показателей

Объемные показатели, как уже говорилось выше, могут быть соизмеримыми (T, pQ, zQ) и несоизмеримыми (объем продукции или товаров разного вида — Q). Соизмеримые объемные показатели могут непосредственно суммироваться и построение сводных индексов этих показателей не вызывает трудностей: они являются обычными относительными величинами и могут быть названы индексами лишь при широком понимании этого термина [см. формулы (10.12) и (10.13)].

Специфические особенности индексного метода используются только при построении сводных индексов несоизмеримых объемных показателей, непосредственное суммирование которых не имеет смысла.

Если отдельные виды единиц совокупности непосредственно несоизмеримы, то для получения их общего итога и построения агрегатного индекса нужно предварительно их соизмерить. Исходя из экономической сущности данного явления нужно найти то общее, что присуще всем видам единиц и использовать эту общую меру в качестве коэффициента соизмерения. Так, различные виды продукции представляют собой разные потребительные стоимости, но все они являются результатами труда и имеют определенную стоимость и ее денежное выражение — цену (p). Каждый продукт имеет также ту или иную себестоимость (z) и трудоемкость (t). Эти качественные показатели и могут быть использованы в качестве коэффициентов соизмерения. Умножая объем продукции каждого вида на соответствующую цену, себестоимость или трудоемкость, мы получим соизмеримые показатели ($pQ, zQ, tQ = T$), которые можно суммировать.

При умножении индексруемого объемного показателя на качественный показатель-соизмеритель каждому виду единиц совокупности придется определенный статистический вес, отражаю-

щий его значение в том или ином экономическом процессе: в процессе образования общей стоимости продукции (ΣpQ), общей ее себестоимости (ΣzQ) и т. п. Таким образом, коэффициенты соизмерения выступают как веса сводного индекса.

Поскольку в индексе объемного показателя в качестве весов могут выступать различные качественные показатели, возникает вопрос о том, какой же именно из них следует использовать. Этот вопрос в каждом конкретном случае должен решаться в соответствии с той познавательной экономической задачей, которая ставится перед индексом, т. е. *выбор тех или иных весов-соизмерителей должен быть обоснован экономически*.

Однако коэффициенты соизмерения носят, как правило, условный характер. Если, например, цена или себестоимость шерстяного костюма вдвое выше, чем хлопчатобумажного, то, конечно, лишь условно можно считать, что при выпуске тысячи шерстяных костюмов вместо тысячи хлопчатобумажных объем продукции увеличится вдвое. В два раза здесь возрастет не количество костюмов, а их стоимость в сопоставимых ценах, т. е. объем продукции в условных единицах измерения.

В практике экономической и статистической работы, когда перед индексами несоизмеримых объемных показателей ставится задача сравнительной характеристики объемов двух совокупностей, т. е. задача обобщения, синтеза, в качестве весов-соизмерителей используются цены. Так, строясь, например, индексы объема промышленной и индексы объема сельскохозяйственной продукции, а также индексы объема товарооборота.

Иначе обстоит дело, если перед индексом ставится задача анализа влияния изменения объемного показателя-фактора на изменение показателя более сложного порядка, т. е. если изменение объема продукции интересует нас не само по себе, а с точки зрения его влияния на изменение общей стоимости продукции, общей ее себестоимости, общих затрат труда и т. п. В этих случаях выбор весов-соизмерителей определяется взаимосвязью анализируемых показателей-факторов, о чем будет сказано ниже.

Чтобы индекс отражал только изменение индексруемого объемного показателя, веса в его числителе и знаменателе должны быть зафиксированы на уровне одного и того же периода.

Если индекс решает задачу сравнения объемов, т. е. рассматривается самостоятельно, вне взаимосвязи с другими индексами, то вряд ли возможно экономически обосновать выбор того периода, на уровне которого фиксируются веса. Трудно, например, сказать, какой индекс будет более правильно характеризовать изменение объема продукции в 1974 г. по сравнению с 1973 г.: с весами базисного 1973 г. ($\frac{\sum Q_{1974}}{\sum Q_{1973}}$) или с весами текущего 1974 г. ($\frac{\sum Q_{1974} p_{1974}}{\sum Q_{1974} p_{1974}}$).

В практике экономической работы веса в индексах динамики объемных показателей обычно фиксируются на уровне базисного периода [см. выше формулу (10.16)].

В отличие от индексов качественных показателей, которые характеризуют соотношение уровней явления применительно к сравнимому кругу единиц (сравнимому ассортименту продукции или товаров), индексы объемных показателей в целях полноты и точности сравнения должны охватывать весь круг единиц, произведенных (или проданных) в соответствующем периоде. Если, например, в базисном периоде производились какие-либо продукты, выпуск которых в текущем периоде был прекращен, а в текущем периоде начали производиться те или иные новые продукты, которых не было в базисном периоде, то в состав продукции каждого периода следует включать все ее виды, фактически производившиеся в данном периоде.

В связи с этим возникает вопрос о том, какие значения весов следует брать для тех видов продукции, которые в одном из периодов не производились. Если, например, в качестве весов берутся цены базисного периода, то как же оценить новые виды продукции, которых в базисном периоде не было?

В практике советской статистики применяются в таких случаях два способа. При расчете индексов объема промышленной продукции новые ее виды, для которых нет цен базисного периода, оцениваются по ценам текущего периода, т. е. для таких продуктов вместо p_0 условно берутся p_1 . При расчете же индексов объема товаров (объема товарооборота в сравнимых ценах) используется метод, основанный на условном предположении, что цены на новые товары изменились в той же степени, что и цены на сравнимый круг товаров. Если, например, цены на сравнимый круг товаров снизились на 10% ($i_p = 0,9$), а цена нового товара в текущем периоде равна 18 руб., то его условная базисная цена составит 20 руб. ($18:0,9$), так что для этого товара $i_p:18:20 = 0,9$.

Оба эти метода не безупречны теоретически, так как носят условный характер. В тех случаях, когда индекс объема рассматривается в системе взаимосвязанных индексов, т. е. используется для оценки степени влияния изменения объемного фактора на изменение функционального показателя, более приемлемым является второй метод.

Ряды индексов с постоянными и переменными весами

При анализе динамики экономических явлений строятся и исчисляются индексы за ряд последовательных периодов. Они образуют ряды либо базисных, либо цепных индексов. В ряду базисных индексов сравнение индексируемого показателя производится с уровнем одного и того же периода, в ряду цепных индексов индексируемый показатель за данный период сопоставляется с уровнем предыдущего периода.

В каждом отдельном индексе веса в числителе и знаменателе обязательно отличаются на одном и том же уровне. Если же строится ряд индексов, то веса в нем могут быть либо постоянными для всех индексов, либо переменными. Если период, на уровне

не которого зафиксированы веса, остается одинаковым для всех индексов ряда, то будем иметь ряд индексов с постоянными весами:

$$\frac{\sum Q_1 p_0}{\sum Q_0 p_0}; \frac{\sum Q_2 p_0}{\sum Q_1 p_0}; \frac{\sum Q_3 p_0}{\sum Q_2 p_0} \text{ и т. д.}$$

Если же период, на уровне которого зафиксированы веса, меняется от индекса к индексу, то получим ряд индексов с переменными весами:

$$\frac{\sum p_1 Q_1}{\sum p_0 Q_1}; \frac{\sum p_2 Q_2}{\sum p_1 Q_2}; \frac{\sum p_3 Q_3}{\sum p_2 Q_3} \text{ и т. д.}$$

Для индексов динамики с постоянными весами сохраняет силу взаимосвязь между цепными и базисными темпами роста:

$$\frac{\sum Q_1 p_0}{\sum Q_0 p_0} \cdot \frac{\sum Q_2 p_0}{\sum Q_1 p_0} = \frac{\sum Q_2 p_0}{\sum Q_0 p_0}.$$

Использование весов какого-либо одного прошлого периода в течение ряда лет дает, таким образом, возможность переходить от цепных индексов к базисным, и наоборот. Поэтому ряды индексов объема продукции и объема товаров строятся в нашей статистике с постоянными весами. Так, в индексах объема промышленной продукции в качестве постоянных весов используются в настоящее время цены, зафиксированные на уровне, который существовал на 1 июля 1967 г. Такие цены, используемые в течение ряда лет, называются сопоставимыми (сравнимыми, неизменными).

Иначе обстоит дело с индексами качественных показателей, которые экономически правильно взвешивать по весам текущего периода. Так как текущий период при переходе от одного индекса к другому меняется, ряды индексов качественных показателей строятся с переменными весами. Переход от цепных индексов к базисным связан в этом случае с определенными погрешностями в расчетах.

Построение территориальных индексов в агрегатной форме

При построении агрегатных территориальных индексов, т. е. при сравнении показателей не во времени, а в пространстве, возникает вопрос о том, на уровне какого района (или объекта) следует зафиксировать веса в числителе и знаменателе индекса.

При изучении динамики, когда за базу сравнения принимается какой-либо прошлый период, индексы качественных показателей, как это было показано выше, более обоснованно взвешивать по весам текущего (сравниваемого) периода. В территориальных же индексах каждый район или объект с одинаковым основанием может быть принят как в качестве сравниваемого, так и в качестве базы сравнения, так что вопрос о том, по весам какого именно района следует взвешивать индекс, остается открытым.

Например, по следующим данным нужно определить, в каком из двух пунктов и на сколько выше уровень цен.

Таблица 10.4
ЦЕНЫ И КОЛИЧЕСТВО ТОВАРОВ, ПРОДАННЫХ НА КОЛХОЗНЫХ РЫНКАХ

Товары	Пункт А		Пункт В	
	продано (т) Q _A	цена 1 кг (руб.) P _A	продано (т) Q _B	цена 1 кг (руб.) P _B
Картофель	400	0,25	600	0,20
Капуста	300	0,16	150	0,25

Если в индексе цен фиксировать веса того пункта, который сравнивается с другим, то можно построить и исчислить два индекса:

$$I_p = \frac{\sum P_A Q_A}{\sum P_B Q_A} = \frac{0,25 \times 400 + 0,16 \times 300}{0,2 \times 400 + 0,25 \times 300} = \frac{148}{135} = 0,955,$$

т. е. применительно к составу товаров, проданных в пункте А, уровень цен в А на 4,5% ниже, чем в В.

Однако, сопоставляя цены пункта В с ценами пункта А, получим:

$$I_p = \frac{\sum P_B Q_B}{\sum P_A Q_B} = \frac{0,2 \times 600 + 0,25 \times 150}{0,25 \times 600 + 0,16 \times 150} = \frac{157,5}{174,0} = 0,905,$$

т. е. применительно к составу товаров, проданных в пункте В, цены в этом пункте ниже, чем в А, на 9,5%.

Таким образом, в каждом пункте уровень цен оказывается более низким, чем в другом, если соотношение уровней цен измерять применительно к кругу товаров сравниваемого пункта.

Чтобы на поставленный вопрос получить вполне определенный, однозначный ответ, возможен различные пути. Достаточно обоснованный и простой путь состоит в том, что в качестве весов принимаются объемы товаров, проданных в обоих районах в целом ($Q = Q_A + Q_B$):

$$I_p = \frac{\sum P_A Q}{\sum P_B Q}.$$

Тогда какой бы район ни был принят за базу сравнения, результаты не будут противоречить друг другу. Так, в нашем примере получим:

$$I_p = \frac{\sum P_A Q}{\sum P_B Q} = \frac{0,25 \times 1000 + 0,16 \times 450}{0,2 \times 1000 + 0,25 \times 450} = \frac{322,0}{312,5} = 1,030;$$

$$I_p = \frac{\sum P_B Q}{\sum P_A Q} = \frac{312,5}{322,0} = 0,970.$$

Следовательно, применительно к кругу товаров, проданных в обоих пунктах в целом, уровень цен в пункте А на 3% выше, чем в В.

Таким образом, при построении территориальных индексов качественных показателей в качестве весов могут быть приняты соответствующие объемные показатели в целом по обоим сравниваемым районам (или объектам). Если же уровни качественных показателей необходимо сравнить с более широкими позициями, например, если имеется в виду сравнивать друг с другом все экономические районы республики, то соответственно нужно расширить и границы той территории, по которой берутся веса.

В территориальных индексах объемных показателей в качестве весов могут быть использованы средние уровни соответствующих качественных показателей (цен, себестоимости, трудоемкости и т. п.), исчисленные в целом по сравниваемым районам:

$$I_Q = \frac{\sum Q_A \bar{P}}{\sum Q_B \bar{P}}, \quad (10.27)$$

где

$$\bar{P} = \frac{P_A Q_A + P_B Q_B}{Q_A + Q_B}.$$

Так, в нашем примере (табл. 10.4) средняя цена картофеля составляет 0,22 руб. $[(0,25 \times 400 + 0,2 \times 600) : 1000]$, а средняя цена капусты — 0,19 руб. $[(0,16 \times 300 + 0,25 \times 150) : 450]$. Используя эти средние цены в качестве весов индекса объема товаров (объема товарооборота в сравнимых ценах), получим:

$$I_Q = \frac{\sum Q_A \bar{P}}{\sum Q_B \bar{P}} = \frac{400 \times 0,22 + 300 \times 0,19}{600 \times 0,22 + 150 \times 0,19} = \frac{145,0}{160,5} = 0,903,$$

т. е. объем продажи товаров (в сопоставимых средних ценах) в пункте А был на 9,7% меньше, чем в пункте В.

Если оценка соотношения объемных показателей производится с более широкими позициями, то и средние уровни, используемые как веса индекса, должны исчисляться по более широкой территории.

Агрегатные индексы — основная форма собственно индексов
Средние индексы

Построение собственно индексов в соответствии с поставленными перед ними экономическими задачами непосредственно приводит, как было показано выше, к индексам в агрегатной форме. В числителе и знаменателе агрегатных индексов находятся суммы произведений индексируемых показателей на связанные с ними показатели — веса, зафиксированные на одном и том же уровне. Экономическое содержание этих сумм произведений в агрегатном индексе выражено непосредственно, что делает отчетливым и ясным экономический смысл как индекса в целом, так и разности между величинами, находящимися в его числителе и знаменателе. Поэтому

му в нашей статистике агрегатная форма рассматривается как основная форма собственно индекса. Она представляет собой исходную базу для построения сводного индекса в другой, производной форме, в форме средней величины из индивидуальных индексов. Вопрос о форме средней и ее весах решается путем преобразования формулы агрегатного индекса. Таким образом, критерием правильности построения среднего индекса является его тождественность агрегатному индексу.

Преобразование агрегатного индекса в средний из индивидуальных индексов производится путем подстановки либо в числитель, либо в знаменатель агрегатного индекса вместо индексированного показателя его выражения через соответствующий индивидуальный индекс. Если такая замена сделана в числителе, то агрегатный индекс окажется преобразованным в средний арифметический, если же в знаменателе — то в средний гармонический из индивидуальных индексов. Практически замена обычно делается там, где в агрегатном индексе находится условная (а не фактическая) величина.

Преобразование агрегатного индекса в средний из индивидуальных возможно только в тех случаях, когда агрегатный индекс построен по сравнимому кругу единиц, для которых могут быть исчислены индивидуальные индексы. Следовательно, в средний из индивидуальных могут быть преобразованы все агрегатные индексы качественных показателей и лишь те индексы объемных показателей, которые охватывают только сравнимый ассортимент продукции или товаров.

Преобразование агрегатного индекса в средний арифметический рассмотрим на примере индекса производительности труда [формула (10.25)]. Так как в этом индексе условные затраты времени ($\Sigma t_i Q_i$) находятся в числителе индекса, то из формулы индивидуального индекса производительности труда, выраженного через трудоемкость единицы продукции ($i_q = t_0 : t_1$), найдем базисную трудоемкость: $t_0 = i_q t_1$. Подставив в числитель агрегатного индекса вместо t_0 выражение $i_q t_1$, получим:

$$I_q = \frac{\Sigma t_0 Q_i}{\Sigma t_1 Q_i} = \frac{\Sigma i_q t_1 Q_i}{\Sigma t_1 Q_i},$$

а так как $t_1 Q_i = T_1$, то

$$I_q = \frac{\Sigma i_q T_1}{\Sigma T_1}. \quad (10.28)$$

Допустим, что при производстве одного вида продукции производительность труда повысилась на 10%, а при производстве другого — на 5%, причем в текущем периоде производством первого вида продукции было занято 30 человек, второго — 20 человек. Поскольку в формуле (10.28) под T могут пониматься как общие затраты времени, так и среднее число работников (или рабочих),

найдем среднее изменение производительности труда по этой формуле:

$$I_q = \frac{\Sigma i_q T_1}{\Sigma T_1} = \frac{110 \times 30 + 105 \times 20}{30 + 20} = 108\%.$$

К расчету сводного индекса как среднего из индивидуальных приходится прибегать в тех случаях, когда отсутствуют данные, необходимые для исчисления индекса в агрегатной форме, но известны индивидуальные или групповые индексы. Так, в государственной розничной торговле учет товарооборота ведется в денежном выражении по группам товаров, данные же о натуральном количестве проданных товаров каждого отдельного вида отсутствуют. В то же время расчетным путем могут быть получены групповые индексы по товарным группам. Это дает возможность исчислять средний индекс цен. Так как при этом числитель агрегатного индекса известен ($\Sigma p_1 Q_1$), то преобразование производится в знаменателе, что приводит к среднему гармоническому индексу: $i_p = p_1 : p_0$, откуда $p_0 = p_1 : i_p$ и

$$I_p = \frac{\Sigma p_1 Q_1}{\Sigma p_0 Q_1} = \frac{\Sigma p_1 Q_1}{\Sigma \frac{p_1 Q_1}{i_p}}. \quad (10.29)$$

Расчет по этой формуле рассмотрим на следующем примере.

Таблица 10.5

ТОВАРООБОРОТ И СНИЖЕНИЕ ЦЕН ПО ГРУППАМ ТОВАРОВ

Группы товаров	Товарооборот за текущий период (тыс. руб.)	Снижение цен в текущем периоде по сравнению с базисным (процентов)	Индекс цен (коэффициент)	Стоимость товаров в текущем периоде по базисным ценам (тыс. руб.)
	$p_1 Q_1$		i_p	$p_0 Q_1 = \frac{p_1 Q_1}{i_p}$
Одежда и белье	4 140	10	0,90	4 600
Трикотаж	2 079	1	0,99	2 100
Чулки и носки	567	19	0,81	700
Итого	6 786	—	—	7 400

$$I_p = \frac{\Sigma p_1 Q_1}{\Sigma \frac{p_1 Q_1}{i_p}} = \frac{6786}{7400} = 0,917, \text{ или } 91,7\%.$$

По всем группам товаров цены снизились в среднем на 8,3%, за счет чего покупатели сэкономили 614 тыс. руб. (6786 — 7400), в том числе при покупке одежды и белья 460 тыс. руб. (4140 — 4600) и т. д.

3. ИНДЕКСНЫЙ МЕТОД АНАЛИЗА ВЛИЯНИЯ ФАКТОРОВ

Социально-экономические явления и процессы связаны между собой, и это находит свое выражение во взаимосвязи соответствующих показателей. Одна из форм взаимосвязи между экономическими показателями состоит в том, что многие из них могут быть выражены в виде произведения нескольких других показателей. Так, общая выработка какого-либо вида продукции может быть представлена как произведение выработки продукции за единицу времени, т. е. уровня производительности труда, на общие затраты времени: $Q = qT$, где $q = Q/T$; товарооборот равен произведению цены товара на его количество и т. п.

Показатели-сомножители во многих подобных случаях выступают как факторы, от величины которых функционально зависит величина показателя-произведения (результативного показателя). Так, увеличение объема продукции является результатом либо увеличения затрат времени, либо повышения производительности труда, либо, наконец, результатом того и другого.

В связи с этим при анализе динамики и выполнения плана возникает задача: выявить и оценить роль отдельных факторов в изменении данного явления, т. е. показать, как изменился результативный показатель за счет изменения каждого отдельного фактора. Решение этой задачи имеет большое практическое значение. С народнохозяйственной точки зрения совсем не безразлично, например, за счет чего и насколько увеличилось производство продукции: за счет ли роста производительности труда или же в результате увеличения числа работников. Рост численности работников имеет определенные пределы, а возможности повышения производительности труда в условиях научно-технической революции поистине безграничны, и именно этот интенсивный фактор должен являться основным источником роста производства. Так, в Директивах XXIV съезда КПСС предусмотрено, что в 1971—1975 гг. 87% общего прироста промышленной продукции должно быть получено за счет роста производительности труда.

Для оценки влияния отдельных факторов, как уже говорилось выше, может быть использован индексный метод. При его использовании для факторного анализа индексы решают аналитическую задачу, выполняют аналитическую функцию.

Основные понятия и условные обозначения
Показатель, изменение которого является результатом изменения других показателей, связанных с ним, называется результативным. Показатели, от которых зависит результативный показатель, называются факторами (факторными показателями, или просто факторами). Индекс результативного показателя, характеризующий его общее изменение за счет изменения всех факторов, называется полным, а индексы, характеризующие степень изменения результативного показателя за счет изменения отдельных факторов, — частными его индексами по соответствующим фак-

торам, или короче, частными индексами этих факторов. Так, если $i_Q = i_T \cdot i_q$, то i_Q — полный, а i_T и i_q — частные индексы результативного показателя Q по факторам T и q (частные индексы этих факторов).

Абсолютный прирост результативного показателя, вызванный изменением всех факторов ($\Delta Q = Q_1 - Q_0$), называется его полным приростом, а прирост, вызванный изменением отдельного фактора, — частным приростом. Частные приросты Q обозначаются через ΔQ с подстрочным значком соответствующего фактора (так, ΔQ_T означает прирост Q за счет T).

Отношение полного или частного прироста к уровню результативного показателя в базисном периоде, называется относительным (полным или частным) приростом:

$$\frac{\Delta Q}{Q_0} = \Delta'Q, \quad \frac{\Delta Q_T}{Q_0} = \Delta'Q_T \text{ и т. п.}$$

Индексный метод факторного анализа позволяет представить полный индекс результативного показателя как произведение частных его индексов. Так, например, $\frac{Q_1}{Q_0} = \frac{T_1}{T_0} \cdot \frac{q_1}{q_0}$, или $i_Q = i_T \cdot i_q$.

Построение подобной системы взаимосвязанных индексов называется геометрическим разложением полного индекса по факторам. Оно характеризует интенсивность изменения результативного показателя, вызванного изменением тех или иных факторов.

Наряду с этим возможно также арифметическое разложение полного прироста и полного относительного прироста по факторам, т. е. разложение их на части (слагаемые), являющиеся результатом изменения отдельных факторов. Арифметическое разложение прироста показывает, насколько в абсолютном выражении увеличился (или уменьшился) результативный показатель в целом и в том числе за счет изменения каждого отдельного фактора. Разложение относительного прироста характеризует то же самое; но в относительном выражении, т. е. в долях или (при умножении их на 100) в процентах к базисному уровню результативного показателя. Так как $\Delta'Q = \frac{\Delta Q}{Q_0} = \frac{Q_1 - Q_0}{Q_0} = i_Q - 1$, то арифметическое разложение относительного прироста — это разложение по факторам разности между полным индексом и единицей (или 100%). Следовательно, арифметическое разложение выражает абсолютную и относительную скорости изменения результативного показателя под влиянием изменения отдельных факторов.

Индексный метод факторного анализа применяется в тех случаях, когда между результативным и факторными показателями существует функциональная связь. Поэтому прежде всего нужно на основе экономического анализа убедиться в наличии функциональной связи между анализируемыми показателями и установить форму этой связи. Форма связи между показателями должна от-

ражать объективную зависимость одного явления от других и реальный характер влияния каждого фактора.

Мы рассмотрим факторный анализ при двух формах связи, которые являются одними из наиболее распространенных в экономике: 1) когда результирующий показатель есть функция произведения двух или нескольких факторных показателей (например, $Q=qT$) и 2) когда результирующий показатель представляет собой сумму произведений показателей-факторов (например, затраты на производство нескольких видов продукции — ΣQ). Будет рассмотрен также частный случай этой формы связи, когда результирующий показатель представляет собой средний уровень качественного показателя.

Построение системы взаимосвязанных индивидуальных индексов

После установления формы связи нужно построить систему взаимосвязанных индексов: представить полный индекс результирующего показателя в виде произведения аналитических частных индексов, т. е. произвести геометрическое разложение полного индекса.

Построение аналитических частных индексов производится с использованием специфических особенностей индексного метода, т. е. взвешивания и элиминирования влияния изменения весов при фиксировании. В каждом частном индексе изменяется только один соответствующий фактор, остальные же факторы фиксируются. В зависимости от порядка фиксирования прочих факторов возможны два метода построения частных индексов при геометрическом разложении полного индекса.

1. Первый метод исходит из того, что все факторы изменяются одновременно и совместно один с другими. Поэтому изменение результирующего показателя является результатом двух процессов: *процесса изолированного, обособленного изменения каждого фактора и процесса их взаимодействия в ходе изменения*. В соответствии с этим задача анализа и состоит прежде всего в том, чтобы выявить влияние обособленного, *изолированного* изменения каждого фактора при условии, что все остальные факторы не изменяются, т. е. сохраняются на уровне базисного периода. Это приводит к фиксированию прочих факторов во всех частных индексах на уровне базисного периода. Так, для взаимосвязи $Q=qT$ будем иметь такую систему взаимосвязанных индексов:

$$\frac{Q_1}{Q_0} = \frac{q_1 T_0}{q_0 T_0} \cdot \frac{q_0 T_1}{q_0 T_0}, \text{ или } i_Q = i_q \cdot i_T. \quad (10.30)$$

¹ В данном случае, т. е. при геометрическом разложении индивидуального индекса, веса в каждом частном индексе, конечно, сокращаются, однако они нужны при последующем переходе к арифметическому разложению прироста Q по факторам q и T .

Если результирующий показатель (y) представляет произведение трех факторов-составителей ($y=wsx$), то система взаимосвязанных индексов будет такой:

$$\frac{y_1}{y_0} = \frac{w_1 s_0 x_0}{w_0 s_0 x_0} \cdot \frac{w_0 s_1 x_0}{w_0 s_0 x_0} \cdot \frac{w_0 s_0 x_1}{w_0 s_0 x_0}. \quad (10.31)$$

Способ элиминирования влияния изменения прочих факторов во всех частных индексах здесь одинаков.

II. Второй метод исходит из того, что сущность, природа факторов и характер их отношения одного к другому неодинаковы: наличие одних факторов есть необходимое условие проявления и возможности измерения других факторов. Так, процесс труда, затраты рабочего времени — условия, необходимые для проявления того или иного уровня производительности труда и, следовательно, для ее измерения. Поэтому, хотя факторы изменяются, как правило, одновременно, исследование влияния их изменения нужно вести *последовательно*, в определенном порядке: *сначала выявить влияние первичного, объемного фактора при фиксировании качественного фактора на уровне базисного периода, а затем — влияние качественного фактора при фиксировании объемного фактора на уровне текущего периода*. Так, для взаимосвязи объема продукции, общих затрат времени и производительности труда ($Q=Tq$) получим такую систему взаимосвязанных индексов:

$$\frac{Q_1}{Q_0} = \frac{T_1 q_0}{T_0 q_0} \cdot \frac{T_0 q_1}{T_0 q_0} \text{ или } i_Q = i_T \cdot i_q. \quad (10.32)$$

В ряде случаев качественный фактор может быть представлен в свою очередь как произведение двух или нескольких субфакторов. Так, выработка продукции за один человеко-день может быть выражена как произведение выработки за один человеко-час на среднюю продолжительность рабочего дня (в часах). Тогда объем продукции будет произведением не двух, а трех факторов:

$$\underbrace{\left[\frac{\text{Объем}}{\text{продукции}} \right]}_y = \underbrace{\left[\frac{\text{Объем продукции}}{\text{Количество затраченных человеко-часов}} \right]}_w \times \underbrace{\left[\frac{\text{Количество затраченных человеко-дней}}{\text{Количество затраченных человеко-дней}} \right]}_s \times \underbrace{\left[\frac{\text{Количество затраченных человеко-дней}}{\text{Количество затраченных человеко-дней}} \right]}_x$$

(результативный показатель) (выработка продукции за 1 человеко-час (качественный фактор)) (средняя продолжительность рабочего дня в часах (качественный фактор)) (объемный фактор)

или $y = wsx$.

В таких и подобных им случаях для построения системы взаимосвязанных индексов вторым методом нужно прежде всего представить результирующий показатель в виде произведения фак-

торов, расположенных в определенной последовательности, вытекающей из их сущности, взаимосвязи и порядка расчета. Эта последовательность должна удовлетворять следующему условию: произведение любого числа смежных факторов должно представлять собой более сложный факторный показатель, имеющий реальный экономический смысл. Так, в нашем примере ws — это выработка продукции за один человеко-день, а zx — количество затраченных человеко-часов.

Первым фактором-сомножителем в этом случае целесообразно взять тот качественный фактор, числителем расчетной формулы которого является результативный показатель (в нашем примере таким фактором является w). Следующим фактором-сомножителем берется тогда тот, числителем расчетной формулы которого является знаменатель первого фактора и т. д. (см. приведенный выше пример). Последним сомножителем будет объемный фактор.

Если результативный показатель является не объемным, как в нашем примере, а качественным, то первым фактором-сомножителем берется тот, числитель которого совпадает с числителем расчетной формулы результативного показателя и который, следовательно, связан непосредственно с результативным показателем. Так, например, если анализируемый результативным показателем является выработка продукции за 1 человеко-день (ws), то первым фактором-сомножителем будет выработка продукции за 1 человеко-час (w), а вторым — средняя продолжительность рабочего дня (s).

Фиксирование весов при построении системы взаимосвязанных индексов вторым методом производится следующим образом. В частном индексе фактора, который является последним в цепи сомножителей, все остальные факторы фиксируются на уровне базисного периода. В каждом следующем частном индексе — если следовать от конца цепи сомножителей к ее началу — уже исследованные (предыдущие) факторы фиксируются на уровне текущего периода, а еще не исследованные — на уровне базисного периода. В этом случае в индексе качественного фактора, который является первым в цепи сомножителей, все веса будут зафиксированы на уровне текущего периода.

В нашем примере такой порядок фиксирования весов приводит к следующей системе взаимосвязанных индексов:

$$\frac{y_1}{y_0} = \frac{w_1 s_1 x_1}{w_0 s_1 x_1} \cdot \frac{w_0 s_1 x_1}{w_0 s_0 x_1} \cdot \frac{w_0 s_0 x_1}{w_0 s_0 x_0} \quad (10.33)$$

Знаменатель каждого индекса-сомножителя служит здесь числителем последующего индекса, как и в цепи факторов. В результате частные индексы являются как бы звеньями единой цепи сомножителей, в связи с чем этот метод получил название цепного метода. Обоснованием такого порядка фиксирования факторов является градация качественных факторов, их ступенчатый переход от низшего порядка к более высокому. Качественные факторы, непосредственно связанные (сопряженные) с объемным,

являются факторами низшего, первого порядка (в примере — s); факторы же, которые связаны с объемным не непосредственно, а через факторы низшего порядка, являются факторами второго и т. д. порядка (в примере — w).

В соответствии с тем, что индексы качественных факторов более обоснованно строить с весами текущего периода, на уровне этого периода и фиксируются все факторы в частном индексе качественного показателя, который в данной цепи имеет самый высокий порядок.

Фиксирование разных факторов на том или ином уровне в отдельных частных индексах обеспечивает возможность перехода от геометрического разложения полного индекса к арифметическому разложению абсолютного и относительного приростов.

Системы взаимосвязанных сводных индексов также могут быть построены двумя методами: методом выявления изолированного влияния изменения факторов и цепным методом. Фиксирование факторов в

частных индексах производится при этом так же, как и при разложении индивидуальных индексов.

Однако при разложении первым методом сводный полный индекс результативного показателя, в отличие от индивидуального индекса, в общем случае не равен произведению частных индексов факторов. Это объясняется тем, что сводные индексы факторов являются средними величинами из индивидуальных индексов, которые обычно не равны один другому. В результате этого в процессе изменения факторов между их индивидуальными индексами могут иметь место эмпирические корреляционные связи. Поэтому полный сводный индекс результативного показателя равен произведению не только частных факторных индексов, но и еще одного индекса-сомножителя — индекса ковариации, т. е. взаимосвязанной вариации факторов в процессе их изменения. Например, сводный индекс затрат на производство (ZzQ) разлагается на произведение не двух, а трех частных индексов:

$$\frac{\sum z_1 Q_1}{\sum z_0 Q_0} = \frac{\sum z_1 Q_1}{\sum z_0 Q_0} \cdot \frac{\sum z_1 Q_1}{\sum z_0 Q_0} \cdot \left(\frac{\sum z_1 Q_1}{\sum z_0 Q_1} \cdot \frac{\sum z_0 Q_0}{\sum z_0 Q_0} \right) \quad (10.34)$$

Первый и второй индексы в правой части этого равенства характеризуют степень влияния обособленного изменения факторов z и Q , т. е. показывают, как изменились бы результативный показатель ($\sum zQ$), если бы z и Q изменялись изолированно. Третий же индекс — индекс ковариации (заклученный в скобки) — характеризует степень дополнительного влияния (сверх влияния изолированного изменения z и Q), которое является результатом того, что факторы z и Q изменялись не изолированно, а взаимосвязанно.

Индекс ковариации можно найти, разделив полный индекс результативного показателя на произведение частных индексов-факторов с весами базисного

периода: $I_{k0z} = I_{zQ} : (I_z \cdot I_Q)$. В теории индексов доказывается, что индекс ковариации может быть выражен так: $I_{k0z} = 1 + r_{IzQ} v_{Iz} v_{IQ}$, где r_{IzQ} — коэффициент корреляции между I_z и I_Q , а v_{Iz} и v_{IQ} — коэффициенты вариации I_z и I_Q .

При использовании *ценного метода* полный индекс разлагается на произведение только частных индексов отдельных факторов. Так, сводный индекс затрат на производство разлагается следующим образом:

$$\frac{\sum z_1 Q_1}{\sum z_0 Q_0} = \frac{\sum z_1 Q_1}{\sum z_1 Q_0} \cdot \frac{\sum z_1 Q_0}{\sum z_0 Q_0} \quad (10.35)$$

Отсутствие третьего индекса, отражающего дополнительное влияние взаимосвязанного изменения факторов, объясняется здесь тем, что дополнительное влияние включено в частный индекс качественного фактора z : этот индекс равен произведению частного индекса, отражающего обособленное влияние фактора z , на индекс ковариации. В случае трех и более факторов дополнительное влияние распределяется между всеми частными индексами качественных факторов в соответствии с более или менее высоким их порядком.

Два рассмотренных метода в общем случае приводят к разным результатам. Неодинаковые результаты получаются и при арифметическом разложении прироста по факторам, которое производится на основе геометрического разложения полного индекса, т. е. на основе системы взаимосвязанных индексов, построенных тем или иным методом. *Более детально разработанным и широко применяемым в практике является ценный метод.* Однако некоторые авторы работ по статистике и экономике считают, что эти методы не противоречат один другому, а дополняют друг друга, углубляя анализ. Дискуссия по этому вопросу продолжается.

Арифметическое разложение прироста по факторам

Рассмотрим сначала выявление *изолированного влияния* факторов. Если мы имеем дело с одним видом единиц совокупности, т. е. с индивидуальными индексами, то полный прирост результативного показателя разлагается на частные приросты, вызванные изолированным изменением соответствующих факторов, и на частный прирост за счет совместного их изменения, которое оказывает дополнительное влияние на полный прирост. Так, если

$$Q = qT, \quad q_1 = q_0 + \Delta q \quad \text{и} \quad T_1 = T_0 + \Delta T,$$

$$\text{то } \Delta Q = Q_1 - Q_0 = q_1 T_1 - q_0 T_0 = (q_0 + \Delta q)(T_0 + \Delta T) - q_0 T_0,$$

откуда

$$\Delta Q = \Delta q T_0 + \Delta T q_0 + \Delta q \Delta T = \Delta Q_q + \Delta Q_T + \Delta Q_{qT}, \quad (10.36)$$

где $\Delta Q_q = \Delta q T_0$ — прирост за счет изолированного изменения q ; $\Delta Q_T = \Delta T q_0$ — прирост за счет изолированного изменения T ; $\Delta Q_{qT} = \Delta q \Delta T$ — прирост за счет того, что оба фактора изменялись не изолированно, а одновременно, совместно.

Графически это можно изобразить так. Отложим на одной оси прямоугольной системы координат базисный и текущий уровни одного фактора-сомножителя, а на другой оси — уровни второго фактора (см. рис. 10.1). Тогда произведение факторов, т. е. результативный показатель, будет изображаться площадью соответствующего прямоугольника. Если оба фактора в текущем периоде увеличились, то три части полного прироста будут изображаться так, как это показано на рис. 10.1.

Сопоставляя разложение полного прироста по факторам с геометрическим разложением Q , т. е. $\frac{Q_1}{Q_0} = \frac{q_1 T_1}{q_0 T_0} = \frac{q_1 T_0}{q_0 T_0} \cdot \frac{q_0 T_1}{q_0 T_0}$, нетрудно убедиться, что *частные приросты за счет отдельных факторов представляют собой разности между числителями и знаменателями соответствующих частных индексов:*

$$\Delta Q_q = \Delta q T_0 = (q_1 - q_0) T_0 = q_1 T_0 - q_0 T_0, \quad (10.37a)$$

$$\Delta Q_T = \Delta T q_0 = (T_1 - T_0) q_0 = T_1 q_0 - T_0 q_0. \quad (10.37b)$$

Это положение справедливо при любом числе факторов. Что же касается прироста за счет совместного изменения всех факторов, то он может быть получен как разность между полным приростом и суммой приростов за счет отдельных факторов, т. е. как остаточная величина:

$$\Delta Q_{qT} = \Delta Q - (\Delta Q_q + \Delta Q_T).$$

Для перехода от разложения абсолютного прироста к разложению относительного прироста достаточно разделить все члены равенства (10.36) на уровень результативного показателя в базисном периоде. После ряда преобразований это приводит к следующей формуле:

$$\frac{\Delta Q}{Q_0} = \Delta' Q = (i_q - 1) + (i_T - 1) + (i_q - 1)(i_T - 1)$$

или

$$\Delta' Q = \Delta' q + \Delta' T + \Delta' q \Delta' T = \Delta' Q_q + \Delta' Q_T + \Delta' Q_{qT}, \quad (10.38)$$

где $\Delta' q = \frac{\Delta q}{q_0}$ и т. д.

Таким образом, *относительный прирост за счет изолированного изменения данного фактора равен относительному приросту самого фактора.*

Это положение также справедливо при любом числе факторов. Относительный же прирост за счет совместного изменения любого числа факторов может быть найден как остаточная величина.

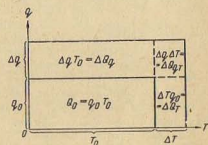


Рис. 10.1. Разложение абсолютного прироста по факторам

Рассмотрим числовой пример. Пусть, объем продукции увеличился с 2000 т (Q_0) до 2600 т (Q_1), а затраты времени возросли со 100 человеко-дней (T_0) до 104 человеко-дней (T_1). Следовательно, уровень производительности труда повысился с 20 т за 1 человеко-день ($q_0 = Q_0 : T_0$) до 25 т (q_1). Используя формулу (10.36), найдем: $\Delta Q = \Delta q T_0 + \Delta T q_0 + \Delta q \Delta T = (25 - 20) \times 100 + (104 - 100) \times 20 + (25 - 20) \times (104 - 100) = 500 + 80 + 20 = 600$ (т).

Следовательно, за счет изолированного изменения производительности труда объем продукции увеличился на 500 т (ΔQ_q), за счет изолированного увеличения затрат времени — на 80 т (ΔQ_T), а в результате совместного изменения обоих факторов — дополнительно еще на 20 т (ΔQ_{qT}). Разделив эти частные приросты на $Q_0 = 2000$ и умножив на 100%, получим разложение относительного прироста: $25\% \left(\frac{500}{2000} \times 100 \right)$, 4% и 1% .

Эти же результаты дает использование формулы (10.38). В целом объем продукции возрос на 30% : $i_Q = Q_1 : Q_0 = 2600 : 2000 = 1,3$, или 130% . Так как производительность труда повысилась при этом на 25% ($i_q = q_1 : q_0 = 25 : 20 = 1,25$), то на столько же процентов по сравнению с базисным уровнем увеличился за счет ее изолированного изменения и объем продукции. Затраты времени возросли на 4% ($i_T = T_1 : T_0 = 104 : 100 = 1,04$), следовательно, за счет их изолированного изменения объем продукции тоже увеличился на 4% . За счет совместного изменения обоих факторов объем продукции возрос на 1% : $\Delta q \Delta T = 0,25 \times 0,04 = 0,01$ (или 1% , или иначе: $30 - (25 + 4) = 1$).

Если результативный показатель охватывает несколько различных видов единиц, выявление изолированного влияния факторов произойдет совершенно аналогично, но на базе взаимосвязанных сводных индексов:

а) абсолютный прирост за счет данного фактора исчисляется как разность между числителем и знаменателем частного сводного индекса этого фактора, а прирост за счет совместного взаимосвязанного изменения факторов — как остаточная величина, т. е. как разность между полным приростом и суммой частных приростов;

б) относительный прирост за счет данного фактора равен относительному приросту самого фактора, а относительный прирост за счет совместного взаимосвязанного изменения факторов равен остаточной величине.

Так, по данным табл. 10.2 (см. с. 359), за счет изолированного роста объема продукции прирост затрат на производство составил: $\Sigma z_0 Q_1 - \Sigma z_0 Q_0 = 100\,000 - 86\,000 = 14\,000$ (руб.). Если бы снижение себестоимости происходило без изменения объема продукции, т. е. при его сохранении на уровне базисного периода, то затраты за счет этого уменьшились бы на 4290 руб. ($\Sigma z_1 Q_0 - \Sigma z_0 Q_0 = 81\,710 - 86\,000$). В целом за счет изолированного изменения обоих факторов затраты возросли бы на 9710 руб. ($14\,000 -$

-4290). Фактически же они увеличились только на 8880 руб. ($\Sigma z_1 Q_1 - \Sigma z_0 Q_0$). Следовательно, за счет совместного взаимосвязанного изменения обоих факторов затраты уменьшились на 830 руб. ($8880 - 9710$).

Эти же результаты можно получить как итоги аналогичных расчетов по отдельным видам продукции.

Таблица 10.6

ВЫЯВЛЕНИЕ ИЗОЛИРОВАННОГО ВЛИЯНИЯ И ВЗАИМОДЕЙСТВИЯ ФАКТОРОВ

Виды продукции	Полный прирост затрат на производство (руб.)	в том числе в результате		
		изолированного изменения		взаимодействия факторов и их процессных изменений
		объема продукции	себестоимости единицы продукции	
	$\Delta Q_z = z_1 Q_1 - z_0 Q_0$	$\Delta Q_{z_0} = z_0 Q_1 - z_0 Q_0$	$\Delta z Q_0 = z_1 Q_0 - z_0 Q_0$	$\Delta Q_{zQ} \text{ или } \Delta(zQ) - (\Delta z Q_0 + \Delta z Q_0)$
Стулья	+6 600	+9 000	-1 800	-600
Стол	+2 000	+4 000	-1 800	-200
Диваны	+280	+1 000	-690	-30
Итого	+8 880	+14 000	-4 290	-830

10 July

Разделив данные итоговой строки этой таблицы на общую сумму затрат на производство в базисном периоде ($\Sigma z_0 Q_0 = 86\,000$), получим разложение относительного прироста по факторам. Так, за счет изолированного изменения себестоимости затраты снизились на 5% [$(-4290 : 86\,000) \times 100$]. Этот же результат можно получить, исчислив относительный прирост самого фактора ($i_z - 1 = \frac{\Sigma z_1 Q_0}{\Sigma z_0 Q_0} - 1$).

Рассмотрим выявление влияния факторов цепным методом. При использовании цепного метода прирост результативного показателя разлагается на столько же частей (слагаемых), на сколько факторов-составляющих разложен результативный показатель, т. е. сколько исследуется факторов. Тот член разложения, который при выявлении изолированного влияния факторов рассматривается как результат совместного взаимосвязанного их изменения, при цепном методе присоединяется к частному приросту за счет качественного фактора, а если качественных факторов несколько, то распределяется между ними.

Частные абсолютные приросты за счет отдельных факторов исчисляются в этом случае как разность между числителем и знаменателем частного индекса, но на базе взаимосвязанных индексов, построенных цепным методом. Это относится как к индивидуальным, так и к сводным индексам при любом числе факторов.

Так, для случая $Q = qT$ на основе формулы (10.32) будем иметь:

$$\Delta Q_T = T_1 q_0 - T_0 q_0 = (T_1 - T_0) q_0, \quad (10.39a)$$

$$\Delta Q_q = T_1 q_1 - T_1 q_0 = (q_1 - q_0) T_1. \quad (10.39b)$$

Таким образом, для одного вида единиц при двух факторах частный абсолютный прирост за счет объемного фактора (T) равен приросту самого объемного фактора, умноженному на базисный уровень качественного фактора (q_0); частный абсолютный прирост за счет качественного фактора равен приросту этого фактора, умноженному на уровень объемного фактора в текущем периоде (T_1). По условиям рассмотренного выше числового примера получим: $\Delta Q_T = (104 - 100) \times 20 = 80(\pi)$; $\Delta Q_q = (25 - 20) \times 104 = 520(\pi)$. Таким образом, из 600 т общего прироста продукции 520 т были получены за счет роста производительности труда — интенсивного фактора. Это составляет 87% общего прироста продукции $[(520 : 600) \times 100]$. Остальные 13% общего прироста были получены за счет экстенсивного фактора — увеличения затрат времени.

По данным табл. 10.2 и по формуле (10.35) определим прирост затрат на производство за счет увеличения объема продукции: $\Sigma z_0 Q_1 - \Sigma z_0 Q_0 = 100\,000 - 86\,000 = 14\,000$ (руб.). За счет же снижения себестоимости единицы продукции затраты уменьшились на 5120 руб.: $\Sigma z_1 Q_1 - \Sigma z_0 Q_1 = 94\,880 - 100\,000 = -5120$. В целом за счет изменения обоих факторов затраты на производство возросли на 8880 руб. $(14\,000 - 5120)$, или $94\,880 - 86\,000$.

Переход к относительным приростам производится путем деления абсолютных приростов на базисный уровень результативного показателя. Так, на основании формулы (10.39) получим:

$$\Delta' Q_T = \frac{\Delta Q_T}{Q_0} = \frac{T_1 q_0 - T_0 q_0}{T_0 q_0} = i_T - 1, \quad (10.40a)$$

$$\Delta' Q_q = \frac{\Delta Q_q}{Q_0} = \frac{T_1 q_1 - T_1 q_0}{T_0 q_0} = i_Q - i_T. \quad (10.40b)$$

Таким образом, при двух факторах относительный прирост за счет объемного фактора равен относительному приросту самого объемного фактора, а относительный прирост за счет качественного фактора равен разности между индексами числителя и знаменателя расчетной формулы этого качественного фактора (расчетная формула $q = Q : T$).

Это положение справедливо не только для двух, но и для любого числа факторов, а также не только для индивидуальных, но и для сводных индексов при условии, что результативный показатель является объемным, а индексы строятся цепным

методом. Так, для затрат на производство по формуле (10.35) будем иметь:

$$\Delta' (\Sigma zQ)_Q = \frac{\Delta (\Sigma zQ)_Q}{\Sigma z_0 Q_0} = \frac{\Sigma z_0 Q_1 - \Sigma z_0 Q_0}{\Sigma z_0 Q_0} = I_Q - 1, \quad (10.41a)$$

$$\Delta' (\Sigma zQ)_z = \frac{\Delta (\Sigma zQ)_z}{\Sigma z_0 Q_0} = \frac{\Sigma z_1 Q_1 - \Sigma z_0 Q_1}{\Sigma z_0 Q_0} = I_{zQ} - I_Q \quad (10.41b)$$

(расчетная формула себестоимости единицы продукции $z = \frac{zQ}{Q}$).

В нашем примере (табл. 10.2) $I_Q = 100\,000 : 86\,000 = 1,163$, или 116,3%. Следовательно, за счет увеличения объема продукции затраты на производство должны были возрасти на 16,3%. А так как фактически они возросли на 10,3% $(94\,880 : 86\,000)$, то снижение себестоимости привело к уменьшению затрат на 6% $(10,3 - 16,3)$.

Наряду с методом выявления изолированного влияния факторов и цепным методом предлагаются также различные варианты третьего метода факторного анализа. Каждый из этих вариантов предусматривает, что дополнительный прирост в результате совместного взаимосвязанного изменения факторов тем или иным путем распределяется между этими факторами: либо поровну, либо пропорционально изолированным приростам и т. п. Однако все эти варианты теоретически небезопасны, имеют условный характер и разработаны применительно лишь к отдельным частным случаям анализа.

Например, для случая, когда каждый из двух факторов в текущем периоде увеличился, академик С. Г. Струминин предлагает такой метод распределения дополнительного прироста, вызванного взаимодействием факторов. Если оба фактора увеличились в одной и той же степени (например, $i_T = i_Q$), то дополнительный прирост (как абсолютный, так и относительный) следует распределить между факторами поровну. Если же один из факторов возрос в меньшей степени, чем другой, то поровну между факторами следует распределить не весь относительный дополнительный прирост, а только ту его часть, которая образовалась бы, если бы второй факторный индекс был равен первому, т. е. меньшему по величине. Эта часть относительного прироста, распределяемая между факторами поровну, равна квадрату разности между факторным индексом и единицей. Остальная часть относительного дополнительного прироста должна быть отнесена к тому фактору, который увеличился в большей степени.

Так, если $i_T = 1,2$, а $i_Q = 1,4$, то относительный дополнительный прирост равен 8% $[(i_T - 1)(i_Q - 1) = 0,2 \times 0,4 = 0,08]$. Если бы оба фактора увеличались

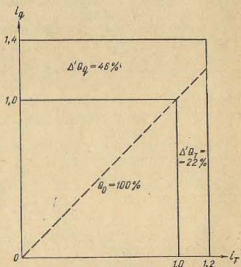


Рис. 10.2. Разложение относительного прироста по факторам методом академика С. Г. Струминина

в 1,2 раза, то дополнительный прирост составил бы 4% ($0,2 \times 0,2$), которые и распределяются поровну между факторами. Остальная часть дополнительного прироста ($8\% - 4\%$) относится к фактору q . В целом, включая также изолированное влияние каждого фактора, частные относительные приросты составят: $\Delta Q_T = 20\% + 2\% = 22\%$; $\Delta Q_q = 40\% + 2\% + 4\% = 46\%$.

При графическом изображении метод, предложенный академиком С. Г. Струмилиным, выглядит так, как показано на рис. 10.2 (применительно к $t_T = 120\%$ и $i_q = 140\%$). Полный относительный прирост делится между факторами прямой, которая является продолжением диагонали квадрата, графически изображающего объем продукции в базисном периоде, принятой за 100%.

Индексный метод анализа динамики среднего уровня качества производственного показателя

На разных участках производства, выпускающих одинаковую продукцию, уровень себестоимости и трудоемкости единицы продукции, уровень производительности труда и других качественных показателей, как правило, неодинаковы. Средний же уровень качественного показателя, исчисленный в целом по всем участкам, зависит как от *уровней на отдельных участках*, так и от *удельных весов этих участков* в общем объеме данной продукции, в общих затратах времени и т. п. Так, средняя себестоимость единицы продукции данного вида зависит от уровней себестоимости на отдельных участках производства и от удельных весов продукции, произведенной на этих участках, в общем ее объеме:

$$\bar{z} = \frac{\sum zQ}{\sum Q} = \sum zq, \quad (10.42)$$

где $q = \frac{Q}{\sum Q}$ — удельный вес единиц продукции, имеющих тот или иной уровень себестоимости, в общем объеме продукции.

Средний уровень производительности труда зависит от ее уровней на отдельных участках и от удельных весов этих участков в общих затратах времени (или в общей численности рабочих):

$$\bar{q} = \frac{\sum Q}{\sum T} = \frac{\sum qT}{\sum T} = \sum qd_T, \quad (10.43)$$

где $d_T = \frac{T}{\sum T}$.

Следовательно, динамика, изменение среднего уровня качественного показателя обусловлены действием двух факторов: 1) изменением уровней этого показателя на отдельных участках (или у отдельных групп единиц совокупности) и 2) изменением удельных весов этих участков (или групп), т. е. изменением структуры совокупности, структурными сдвигами.

В связи с этим при анализе динамики среднего уровня возникает вопрос о том, в какой мере изменение среднего уровня было вызвано действием каждого из этих двух факторов в отдельности. Например, на сколько рублей и на сколько процентов изменилась средняя себестоимость единицы продукции за счет

изменения уровня себестоимости на отдельных участках и на сколько за счет изменения удельных весов участков в общем объеме продукции, т. е. за счет сдвигов в структуре совокупного объема продукции, сдвигов в структуре производства продукции.

Ответы на подобные вопросы имеют большое познавательное и практическое значение. В наших планах для повышения эффективности производства предусматривается использование каждого из этих двух факторов. Известно, например, что добыча угля открытым способом (в карьерах) обходится значительно дешевле, чем подземным (в шахтах), а производительность труда при этом в несколько раз выше, чем на шахтах. Поэтому добычу угля открытым способом в девятой пятилетке предусмотрено развивать опережающими темпами, чтобы в 1975 г. довести удельный вес добычи этим способом в общей добыче до 30%.

Анализ динамики среднего уровня аналогичен анализу изменения результативного показателя, представляющего сумму произведений двух факторов. Однако вместо значений качественного и объемного факторов у единиц разного вида мы имеем здесь дело с уровнями качественного показателя у отдельных групп единиц одного и того же вида, но относящихся к разным участкам производства, и с их удельными весами в общей численности единиц. Аналогичны и методы анализа в этих двух случаях.

Анализ динамики среднего уровня производится путем построения системы взаимосвязанных индексов. При этом индекс среднего уровня, характеризующий его динамику за счет изменения обоих факторов, разлагается на произведение аналитических индексов-сомножителей, каждый из которых отражает изменение только одного фактора и тем самым влияние этого изменения на динамику среднего уровня.

Система взаимосвязанных индексов может быть построена либо методом выявления *изолированного влияния* изменения факторов, либо *цепным методом*.

При выявлении *изолированного влияния* изменения факторов система взаимосвязанных индексов средней себестоимости выглядела так:

$$\frac{\sum z_1 d_1}{\sum z_0 d_0} = \frac{\sum z_1 d_0}{\sum z_0 d_0} \cdot \frac{\sum z_1 d_1}{\sum z_1 d_0} \cdot \left(\frac{\sum z_1 d_1}{\sum z_1 d_0} \cdot \frac{\sum z_1 d_0}{\sum z_1 d_0} \right), \quad (10.44)$$

где $d_q = \frac{Q}{\sum Q}$ — удельные веса участков в общем объеме продукции (подстрочный значок u в формуле опущен).

Индекс в левой части этого равенства представляет отношение средней себестоимости в текущем периоде к средней себестоимости в базисном периоде, т. е. индекс средней себестоимости (i_z). Он характеризует ее изменение в целом, за счет изменения всех факторов, и называется *индексом переменного состава*, так как удельные веса (d), выражающие состав, т. е. структуру, в его числителе и знаменателе взяты неодинаковые.

Первый индекс-сомножитель в правой части равенства показы-

вает, как изменилась бы средняя себестоимость только за счет изменения ее уровней на отдельных участках при сохранении удельных весов участков в общем объеме продукции на уровне базисного периода. Этот индекс называется индексом постоянного (факсированного) состава и может быть представлен как среднеарифметическая взвешенная из индексов на отдельных участках (\bar{i}).

Второй индекс-сомножитель показывает, как изменилась бы средняя себестоимость только в результате структурных сдвигов, т. е. только за счет изменения удельных весов участков в общем объеме продукции при сохранении себестоимости единицы продукции на всех участках на базисном уровне. Этот индекс называется индексом влияния структурных сдвигов на изменение средней себестоимости, или, короче, индексом структурных сдвигов ($i_{стр. сдв}$).

Наконец, последний индекс-сомножитель (заклученный в скобки) показывает, как изменилась бы средняя себестоимость в результате совместного взаимосвязанного изменения себестоимости на отдельных участках и их удельных весов в общем объеме продукции. Это индекс ковариации ($i_{ков}$).

Рассмотрим расчет этих индексов на следующем примере.

Таблица 10.7

ПРОИЗВОДСТВО И СЕБЕСТОИМОСТЬ МОЛОКА

Типы хозяйств	Базисный период			Текущий период		
	производство молока		себестоимость 1 т (руб.)	производство молока		себестоимость 1 т (руб.)
	тыс. т	доля итога		тыс. т	доля итога	
	Q_0	d_0	z_0	Q_1	d_1	z_1
Специализированные	10	0,1	155	22	0,2	140
Неспециализированные	90	0,9	205	88	0,8	192
Итого	100	1,0	200	110	1,0	181,6

В базисном периоде специализированные хозяйства, где себестоимость значительно ниже, произвели только 10% всего молока, а в текущем периоде — 20%. Следовательно, изменилась структура общего производства молока. Изменилась и себестоимость 1 т молока в каждом типе хозяйств: в специализированных хозяйствах она снизилась на 15 руб., или на 9,7%, в неспециализированных — на 13 руб., или на 6,3%.

Средняя себестоимость 1 т в базисном периоде составляла 200 руб. ($\Sigma z_0 d_0 = 155 \times 0,1 + 205 \times 0,9$). Если бы структура производства оставалась такой, какой она была в базисном периоде, а себестоимость 1 т в каждом типе хозяйств снизилась до уровня текущего периода, то в результате такого изолированного сниже-

ния себестоимости ее средний уровень понизился бы до 186,8 руб. ($\Sigma z_1 d_0 = 140 \times 0,1 + 192 \times 0,9$), т. е. на 13,2 руб., или на 6,6% ($\Sigma z_1 d_0 : \Sigma z_0 d_0 = 0,934$).

Наоборот, если бы в каждом типе хозяйств себестоимость 1 т оставалась на уровне базисного периода, а изменилась бы только структура производства, т. е. только удельные веса хозяйств разного типа в общем производстве молока, то за счет такого изолированного изменения структуры средняя себестоимость понизилась бы до 195 руб. ($\Sigma z_0 d_1 = 155 \times 0,2 + 205 \times 0,8$), или на 2,5% по сравнению с базисным периодом ($\Sigma z_0 d_1 : \Sigma z_0 d_0 = 0,975$).

В целом за счет обоих этих изолированных изменений средняя себестоимость должна была бы снизиться на 8,9% ($0,934 \times 0,975 = 0,9106$). Однако фактически эти изменения происходили не изолированно, а одновременно и взаимосвязанно. В результате средняя себестоимость в текущем периоде снизилась до 181,6 руб. ($\Sigma z_1 d_1 = 140 \times 0,2 + 192 \times 0,8$), т. е. уменьшилась по сравнению с базисным периодом на 18,4 руб., или на 9,2% ($\Sigma z_1 d_1 : \Sigma z_0 d_0 = 0,908$). Следовательно, за счет совместного взаимосвязанного изменения себестоимости в каждом типе хозяйств и изменения структуры производства средняя себестоимость дополнительно понизилась на 0,3% [$0,908 : (0,934 \times 0,975) = 0,997$]. В абсолютном выражении это дополнительное снижение составило 0,2 руб. [$184 - (13,2 + 5,0)$].

Система взаимосвязанных индексов среднего уровня производительности труда строится аналогично:

$$\frac{\Sigma q_1 d_1}{\Sigma q_0 d_0} = \frac{\Sigma q_1 d_0}{\Sigma q_0 d_0} \cdot \frac{\Sigma q_0 d_1}{\Sigma q_0 d_0} \cdot i_{ков}, \quad (10.45)$$

где $d_T = \frac{T}{\Sigma T}$ — удельные веса участков в общих затратах времени или в общей численности рабочих (подстрочный значок u в формуле опущен).

Индекс среднего уровня производительности труда (индекс переменного состава) может быть исчислен и по другим тождественным формулам:

$$i_q = \frac{\Sigma q_1 d_1}{\Sigma q_0 d_0} = \frac{\Sigma q_1 T_1}{\Sigma T_1} : \frac{\Sigma q_0 T_0}{\Sigma T_0} = \frac{\Sigma Q_1}{\Sigma T_1} : \frac{\Sigma Q_0}{\Sigma T_0} = \frac{\bar{q}_0}{\bar{q}_1}. \quad (10.46)$$

Этот индекс характеризует изменение среднего уровня производительности труда в целом, т. е. за счет всех факторов.

Индекс постоянного состава ($\bar{i}_q = \Sigma q_1 d_0 : \Sigma q_0 d_0$) показывает, как изменился бы средний уровень производительности труда только за счет изменения ее уровней на отдельных участках при сохранении удельных весов участков в общих затратах времени на уровне базисного периода.

Индекс влияния структурных сдвигов на изменение среднего уровня производительности труда ($i_{стр. сдв} = \Sigma q_0 d_1 : \Sigma q_0 d_0$) характеризует влияние изолированного изменения структуры на \bar{q} .

Индекс ковариации характеризует степень влияния совместного взаимосвязанного изменения обоих указанных факторов на изменение средней производительности труда (\bar{q}).

При цепном методе анализа динамики среднего уровня индекс переменного состава разлагается на два индекса: индекс постоянного состава и индекс влияния структурных сдвигов на изменение среднего уровня. При этом используются две различные системы взвешивания и две схемы разложения индекса переменного состава. В одной из них индекс постоянного состава, как индекс качественного показателя, строится с весами текущего периода, а индекс влияния структурных сдвигов — с весами базисного периода, в результате чего влияние ковариации оказывается отнесенным к качественному фактору (например, к z):

$$\frac{\sum z_i d_i}{\sum z_i d_0} = \frac{\sum z_i d_i}{\sum z_i d_i} \cdot \frac{\sum z_i d_i}{\sum z_i d_0}, \quad (10.47)$$

где $d = \frac{Q}{\Sigma Q}$.

В нашем примере (см. табл. 10.7) получим:

$$i_{\text{стр. сдв}} = \frac{\sum z_i d_i}{\sum z_i d_0} = \frac{195}{200} = 0,975,$$

т. е. за счет структурных сдвигов средняя себестоимость снизилась на 5 руб., или на 2,5% по сравнению с базисным уровнем;

$$\bar{i}_z = \frac{\sum z_i d_i}{\sum z_i d_i} = \frac{181,6}{195,0} = 0,931,$$

т. е. за счет снижения себестоимости на отдельных участках средняя себестоимость снизилась еще на 13,4 руб., или на 6,9% по сравнению с условным ее уровнем ($\Sigma z_0 d_i$).

Во второй схеме разложения индекса постоянного состава строится с весами базисного периода ($\bar{i}_z = \Sigma z_i d_0 : \Sigma z_0 d_0$), а индекс структурных сдвигов — с весами текущего периода ($i_{\text{стр. сдв}} = \Sigma z_i d_i : \Sigma z_i d_0$), в результате чего влияние ковариации относится к влиянию структурных сдвигов. По условиям нашего примера индекс структурных сдвигов оказывается равным 0,972.

Аналогично строится и система взаимосвязанных индексов среднего уровня производительности труда. Так, при взвешивании индекса постоянного состава по весам текущего периода будем иметь:

$$\frac{\sum q_i d_{T_i}}{\sum q_i d_{T_0}} = \frac{\sum q_i d_{T_i}}{\sum q_i d_{T_i}} \cdot \frac{\sum q_i d_{T_i}}{\sum q_i d_{T_0}}, \quad (10.48)$$

где $d_T = T : \Sigma T$ — удельные веса участков в общих затратах времени.

Если вместо прямого показателя производительности труда

($q = Q : T$) использовать обратный ($t = T : Q$), то система индексов будет такова:

$$\frac{\sum t_i d_{Q_i}}{\sum t_i d_{Q_0}} = \frac{\sum t_i d_{Q_i}}{\sum t_i d_{Q_i}} \cdot \frac{\sum t_i d_{Q_i}}{\sum t_i d_{Q_0}}, \quad (10.49)$$

где $d_Q = Q : \Sigma Q$ — удельные веса участков в общем объеме продукции.

Индексы постоянного состава, а соответственно и индексы структурных сдвигов в формулах (10.48) и (10.49) имеют различный экономический смысл. Индекс постоянного состава из формулы (10.48) можно преобразовать так:

$$\frac{\sum q_i d_{Q_i}}{\sum q_0 d_{Q_i}} = \frac{\sum q_i T_i}{\sum q_0 T_i} = \frac{\sum Q_i}{\sum Q_0}.$$

Здесь продукция текущего периода (ΣQ_i) сопоставляется с той продукцией, которая была бы произведена, если бы на отдельных участках производительность труда сохранилась на уровне базисного периода, а затраты времени были такими, как в текущем периоде ($\Sigma q_0 T_i$). Следовательно, изменение производительности труда оценивается здесь с точки зрения его влияния на изменение объема продукции при фиксированных затратах времени на отдельных участках.

Индекс постоянного состава из формулы (10.49) может быть выражен таким образом:

$$\frac{\sum t_i d_{Q_i}}{\sum t_i d_{Q_0}} = \frac{\sum t_0 Q_i}{\sum t_i Q_i} = \frac{\sum T_0}{\sum T_i}.$$

т. е. здесь изменение производительности труда оценивается в другом аспекте — с точки зрения его влияния на изменение общих затрат времени при фиксированном объеме продукции на отдельных участках.

Соответственно и индексы структурных сдвигов характеризуют изменение среднего уровня производительности труда за счет изменения в структуре разных явлений: в первом случае (формула 10.48) — в структуре общих затрат времени (или численности рабочих), а во втором (формула 10.49) — в структуре производств продукции.

Выбор той или другой системы определяется конкретными экономическими задачами анализа.

4. ИНДЕКСЫ В БУРЖУАЗНОЙ СТАТИСТИКЕ

Советская статистика при построении индексов исходит из экономической сущности изучаемых явлений и тех конкретных познавательных задач, которые должны быть решены с помощью этих индексов. Каждый индекс в нашей статистике имеет экономическое обоснование и определенный экономический смысл.

Иначе подходит к построению индексов *буржуазная статистика*, рассматривающая эту проблему с *формально-математических и субъективистских позиций*. Так, американский статистик И. Фишер в своей работе «Построение индексов» прямо пишет, что «...нашей задачей является найти такую математическую формулу, которая совершенно не зависела бы от природы того материала, к которому мы ее будем применять»¹.

При построении сводных индексов Фишер исходит из того, что они должны обладать теми же математическими свойствами, что и индивидуальные индексы. В соответствии с этим Фишер выдвигает ряд формальных требований, которым должна удовлетворять *«идеальная»* формула индекса. Обнаружив, что агрегатные индексы цен, взвешенные один по отчетным, а другой по базисным весам, этим требованиям не удовлетворяют, Фишер *«скрещивает»* их между собой, вычисляя из них *среднюю геометрическую*:

$$\sqrt{\frac{\sum p_1 Q_0}{\sum p_0 Q_0} \cdot \frac{\sum p_1 Q_1}{\sum p_0 Q_1}}$$
 Эту формулу, лишенную конкретного экономического смысла, Фишер и считает «идеальной», так как только она удовлетворяет тем формально-математическим требованиям, которые Фишер считает основными, главными. Однако другие буржуазные теоретики считают основными иные требования и приходят к другим индексным формулам. Различный подход к выводу тех основных требований, которым должен, по их мнению, удовлетворить индекс, придает буржуазным теориям индексов субъективистский характер.

Формальный характер носят и попытки построить систему взаимосвязанных аналитических индексов. Так, И. Монтомери разлагает индекс стоимости товаров на факторные индексы цен и объема товаров следующим образом:

$$\frac{\sum p_1 Q_1}{\sum p_0 Q_0} = \left(\frac{\sum p_1 Q_1}{\sum p_0 Q_1} \right)^a \cdot \left(\frac{\sum p_1 Q_0}{\sum p_0 Q_0} \right)^{1-a}$$

Факторные индексы выражаются здесь как функции результативного показателя, и хотя математическое тождество здесь несомненно, столь же несомненно, что величина a имеет здесь лишь формально-математическое содержание и значение.

Отсутствие экономического обоснования и голый формализм при построении индексов сочетаются в буржуазной статистике с существенными дефектами в практике их исчисления. Прежде всего буржуазная статистика обычно не располагает полными и точными данными, необходимыми для исчисления большинства индексов. Так, например, если в нашей статистике индекс объема продукции вычисляется на основе прямой оценки в сопоставимых ценах всей выпущенной продукции, то буржуазная статистика вынуждена исчислять этот индекс по неполным, частичным данным, которые охватывают не всю продукцию, а лишь определенный набор *«продуктов-представителей»*. Кроме того, в связи с отсутствием данных о производстве тех или иных *«продуктов-предста-*

вителей» в натуральном выражении буржуазная статистика использует вместо них всякого рода *косвенные* показатели, например, данные о потреблении сырья, об отработанном времени и т. п. Не по всему кругу товаров, а лишь по условному их набору исчисляется и индекс цен.

Различные условности, к которым прибегает буржуазная статистика при исчислении индексов, открывают широкие возможности для прямой их *фальсификации* в тех случаях, когда это связано с классовыми интересами монополистической буржуазии. Особенно беззастенчиво фальсифицирует буржуазная статистика индексы реальной заработной платы и индексы стоимости жизни (индексы «потребительских цен», охватывающие цены на товары и услуги). Дело в том, что в условиях роста дороговизны трудящимся в ряде случаев удается добиться так называемой «скользящей шкалы» заработной платы, в соответствии с которой ставки заработной платы должны автоматически повышаться при вздорожании жизни. В связи с этим официальная статистика в угоду буржуазии стремится всеми способами занижать индексы стоимости жизни. Обычно эти индексы исчисляются как среднеарифметические взвешенные из индивидуальных индексов цен *«репрезентативного»* набора товаров и услуг. Веса товаров и услуг определяются на основе бюджетных обследований семей рабочих и служащих, из числа которых исключаются, однако, семьи низкооплачиваемых рабочих и семьи безработных. Это нарушает репрезентативность набора и искажает веса: товары первой необходимости и массового потребления, цены на которые растут особенно быстро, либо попадают в набор с преуменьшенными весами, либо совсем исчезают из него. В результате индексы стоимости жизни оказываются заниженными.

Искажая и фальсифицируя индексы, буржуазная статистика пытается скрыть усиление эксплуатации трудящихся и другие пороки и язвы капитализма. В то же время буржуазные экономисты всячески стремятся преуменьшить успехи стран мировой социалистической системы, подвергая сомнению и пересчет публикуемые в СССР и других социалистических странах индексы роста объема производства. В связи с этим американский экономист Кемпбелл писал: «Если кто-либо достаточно сильно хочет поверить, что советское промышленное производство не росло слишком быстро, то всегда можно найти такую систему взвешивания (индексов. — Авт.), которая даст это заключение»¹. Однако к какой бы системе взвешивания ни прибегали апологеты империализма при исчислении индексов, превосходство социализма над капитализмом в области темпов развития экономики остается реальным фактом.

¹ Цитируется по журналу: «Мировая экономика и международные отношения», 1962, № 1, с. 145.

¹ Фишер И. Построение индексов, М., 1928, с. 181.

ОГЛАВЛЕНИЕ

От авторов	3
Глава I. Статистика как общественная наука	5
1. Возникновение и развитие статистики	5
2. Предмет статистики	9
3. Теоретические основы статистики	17
4. Метод статистики	19
5. Статистика как наука	20
Глава II. Современная организация статистики	25
1. Учет и статистика в социалистическом обществе	25
2. Организация статистики в СССР	27
3. Организация статистики в зарубежных социалистических странах	36
4. Организация статистики в капиталистических странах	38
5. Международные статистические организации	41
Глава III. Статистическое наблюдение	43
1. Статистическое наблюдение и его задачи	43
2. Формы, виды и способы статистического наблюдения	44
3. План статистического наблюдения	49
4. Практика проведения статистического наблюдения в СССР	58
5. Строгая достоверность и тщательная проверка статистических данных—закон статистического наблюдения в СССР	66
Глава IV. Сводка и группировка материалов статистического наблюдения	75
1. Содержание и задачи статистической сводки	75
2. Статистические группировки	78
3. Статистические таблицы	98
4. Графические способы изображения статистических данных	104
5. Статистические ряды	111
6. Организация и техника сводки	117
Глава V. Абсолютные и относительные величины	123
1. Абсолютные величины	123
2. Относительные величины	125
3. Общие принципы построения и использования абсолютных и относительных статистических величин	134
4. Графические способы изображения абсолютных и относительных величин	139
Глава VI. Средние величины и показатели вариации	147
1. Сущность и значение средних величин	147
2. Способы вычисления средних	151
3. Медиана и мода	170
4. Показатели вариации	175
5. Анализ вариационных рядов	184
6. Критика буржуазных «теорий» средних и приемов их вычисления	192

Глава VII. Выборочное наблюдение	196
1. Основные положения выборочного метода	196
2. Способы отбора, обеспечивающие репрезентативность выборки	199
3. Малая выборка	217
4. Моментно-выборочное наблюдение	220
5. Распространение выборочных данных на генеральную совокупность	222
6. Практика применения выборочного наблюдения	223
Глава VIII. Статистическое изучение взаимосвязей	228
1. Взаимосвязи общественных явлений и задачи их статистического изучения	228
2. Основные статистические методы и приемы изучения стохастических взаимосвязей	232
3. Анализ регрессий и корреляций (регрессионно-корреляционный анализ)	245
Глава IX. Ряды динамики	292
1. Ряды динамики и их виды	292
2. Показатели анализа динамики	295
3. Средние показатели динамики	301
4. Способы графического изображения динамики	315
5. Основные приемы обработки и анализа рядов динамики	319
Глава X. Индексы	346
1. Общее понятие об индексах и индексном методе	346
2. Сводные индексы сравнения уровней социально-экономических явлений	356
3. Индексный метод анализа влияния факторов	370
4. Индексы в буржуазной статистике	387

728, 723, 725, 747, 759, 770, 775, 292, 295,
301-346, 356, 370.

ОГЛ

Гла

Гла

Гла

Гла

Гла

Гла

Владимир Сергеевич Козлов
Павел Павлович Полушин
Федот Григорьевич Долгушевский
Яков Моисеевич Зраих

ОБЩАЯ ТЕОРИЯ СТАТИСТИКИ

Редактор Л. И. Ларина
Техн. редактор Р. Н. Феоктистова
Корректор Т. М. Васильева
Худ. редактор Т. В. Стихно
Переплет художника И. И. Карпикова
Сдано в набор 21/XI 1974 г.
Подписано к печати 30/IV 1975 г.
Формат бумаги 60 × 90/16. Бумага № 2
Объем 24,5 печ. л. Уч.-изд. л. 27,68
Тираж 30 000 экз. А. 03119
(Тематич. план 1975 г. № 105)

Издательство «Статистика», Москва, ул. Кирова, 39.
Заказ № 62. Цена 1 р. 13 к.

Типографии им. Котлякова издательства «Финансы»
Государственного комитета Совета Министров СССР по
делам издательства, полиграфии и книжной торговли,
191023, Ленинград, Д-23, Садовая, 21.